

\*\*\*\*\*

W P E R L H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 06:17:41 1998; MasPar time 326.16 Seconds  
Tabular output not generated. 543.181 Million cell updates/sec

Title: >US-08-951-733-19  
Description: (1-3798) from US08951733.seq  
Perfect Score: 3798  
N.A. Sequence: 1 CCACGGGTCGGCGAGCGCT.....GGATAGTCATCCCTGAT 3798  
Comp: GGTGCGAGCGCGTCCGGA.....CCTTATCAGTAGGCGACTA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 9.574; Variance 5.701; scale 1.679

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	ID	Description	Pred. No.
1	67	1.8	7218	1	US-08-232- Sequence 14, Applicati	1.46e-25
2	49	1.3	7218	1	US-08-232- Sequence 14, Applicati	1.85e-14
3	34	0.9	215	1	US-08-238- Sequence 5, Applicatio	7.21e-06
4	35	0.9	215	1	US-08-238- Sequence 5, Applicatio	2.06e-06
5	25	0.7	66	1	US-08-471- Sequence 144, Applicat	2.91e-01
6	25	0.7	68	1	US-07-977- Sequence 243, Applicat	2.91e-01
7	25	0.7	69	1	US-08-471- Sequence 142, Applicat	2.91e-01
8	26	0.7	74	2	PCT-US95-1 Sequence 94, Applicati	2.91e-01
9	25	0.7	74	2	PCT-US95-1 Sequence 100, Applicati	2.91e-01
10	26	0.7	74	2	PCT-US95-1 Sequence 100, Applicati	2.91e-01
11	26	0.7	74	2	PCT-US95-1 Sequence 99, Applicati	2.91e-01
12	25	0.7	75	2	PCT-US95-1 Sequence 99, Applicati	2.91e-01
13	25	0.7	75	2	PCT-US95-1 Sequence 99, Applicati	2.91e-01
14	26	0.7	81	2	PCT-US95-1 Sequence 92, Applicati	2.91e-01
15	25	0.7	81	2	PCT-US95-1 Sequence 98, Applicati	2.91e-01
16	26	0.7	81	2	PCT-US95-1 Sequence 98, Applicati	2.91e-01
17	26	0.7	81	2	PCT-US95-1 Sequence 92, Applicati	2.91e-01
18	26	0.7	82	2	PCT-US95-1 Sequence 97, Applicati	2.91e-01
19	26	0.7	82	2	PCT-US95-1 Sequence 97, Applicati	2.91e-01

c	20	25	0.7	242	1	US-08-273- Sequence 1, Applicatio	2.91e-01
c	21	25	0.7	1004	2	PCT-US95-0 Sequence 7, Applicatio	2.91e-01
c	22	25	0.7	1386	2	PCT-US95-0 Sequence 2, Applicatio	2.91e-01
c	23	25	0.7	1611	2	PCT-US93-0 Sequence 3, Applicatio	2.91e-01
c	24	23	0.6	65	1	US-08-471- Sequence 145, Applicat	2.49e+00
c	25	24	0.6	65	1	US-08-471- Sequence 145, Applicat	8.62e-01
c	26	24	0.6	66	2	PCT-US95-1 Sequence 93, Applicati	8.62e-01
c	27	24	0.6	66	1	US-08-471- Sequence 144, Applicat	8.62e-01
c	28	23	0.6	68	1	US-08-471- Sequence 143, Applicat	8.62e-01
c	29	24	0.6	68	1	US-08-471- Sequence 143, Applicat	8.62e-01
c	30	24	0.6	69	1	US-08-471- Sequence 142, Applicat	8.62e-01
c	31	23	0.6	84	1	US-08-300- Sequence 25, Applicati	2.49e+00
c	32	23	0.6	84	2	PCT-US94-0 Sequence 25, Applicati	2.49e+00
c	33	23	0.6	84	1	US-08-209- Sequence 23, Applicati	2.49e+00
c	34	23	0.6	84	1	US-08-133- Sequence 120, Applicat	2.49e+00
c	35	23	0.6	108	1	US-08-357- Sequence 2, Applicatio	2.49e+00
c	36	23	0.6	225	1	US-08-299- Sequence 18, Applicati	2.49e+00
c	37	23	0.6	225	1	US-07-807- Sequence 15, Applicati	2.49e+00
c	38	24	0.6	242	1	US-08-273- Sequence 1, Applicatio	8.62e-01
c	39	23	0.6	1404	3	5480796-8 Patent No. 5480796.	2.49e+00
c	40	23	0.6	1640	1	US-07-807- Sequence 11, Applicati	2.49e+00
c	41	23	0.6	1640	1	US-08-299- Sequence 11, Applicati	2.49e+00
c	42	23	0.6	2253	3	5457037-2 Patent No. 5457037.	2.49e+00
c	43	24	0.6	2454	1	US-07-872- Sequence 1, Applicatio	8.62e-01
c	44	23	0.6	2625	3	5457037-4 Patent No. 5457037.	2.49e+00
c	45	23	0.6	3336	3	5457037-1 Patent No. 5457037.	2.49e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx  
DT  
DE Sequence 14, Application US/08232463  
CC Sequence 14, Application US/08232463  
CC Patent No. 5670367  
CC GENERAL INFORMATION:  
CC APPLICANT: DORNER, F.  
CC APPLICANT: SCHEIFLINGER, F.  
CC APPLICANT: FALKNER, F. G.  
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 1800 Diagonal Road, Suite 500  
CC CITY: Alexandria  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22313-0299  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/232,463  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/995,313  
CC FILING DATE:  
CC APPLICATION NUMBER: EP 91 114 300.6  
CC FILING DATE: 26-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703)836-9300  
CC TELEFAX: (703)833-4109  
CC TELEX: 899149  
CC INFORMATION FOR SEQ ID NO: 14:





[illegible]

```
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc.feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 0.9%; Score 35; DB 1; Length 215;
Best Local Similarity 13.2%; Pred.No. 2.06e-06;
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSSVSVRTASCNDRKAKDKGNTTSWTTDDCCNRTWGVCDDTITYRVNNDSGHNKYSSAN 65
Cp 1616 AGTCCCAGCGCTCATCTCCACGTCTCTGCAGCGAGACTTGCGATCTTCCTCCCA 1557
Db 66 YNTGGNNVGAATHYTHYNVSADSKTVTDYSNASGTSSTSSNGTGDGNRSGADSYGSST 125
Cp 1556 GGGAGATGAACCTCTTGGTGTTCTCTGAGGAAGCGGCTTCGTTCCTGGAGCCCCAGA 1497
Db 126 AMTSRRNRKTANNVDSRNMGDASVSDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGT 185
Cp 1496 GGCTCGGGGCCACCGCGCGCAGCGAGCGCCGCGAAGCGGTACACCTGCCAGGGGC 1437
Db 186 KSNVSNNGGKKRDYS 202
Cp 1436 TGCTGTCTGGCGGAGC 1420

RESULT 5
ID US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.
AC xxxxxx
DT
DE Sequence 144, Application US/08471052A
CC Sequence 144, Application US/08471052A
CC Patent No. 5625033
CC GENERAL INFORMATION:
CC APPLICANT: Kay, B. K.
CC APPLICANT: Fowlkes, D. M.
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/471,052A
CC FILING DATE: 06-JUNE-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
```

CC	LENGTH: 68
CC	TYPE: NUCLEIC ACID
CC	STRANDEDNESS: SINGLE
CC	TOPOLOGY: LINEAR
CC	ANTI-SENSE: NO
SQ	SEQUENCE 68 BP; 7 A; 38 C; 17 G; 6 T; 0 OTHER.
Query Match      0.7%; Score 25; DB 1; Length 68;	
Best Local Similarity 75.5%; Pred.No. 2.91e-01;	
Matches	37; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Db	11 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGGACCTTCCTTGAGG 59 
Qy	254 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGGTGCTGCCTGAAGG 302
RESULT 7	
ID	US-08-471-052A-142 STANDARD; DNA; UNC; 69 BP.
AC	xxxxxx
DT	
DE	Sequence 142, Application US/08471052A
CC	Sequence 142, Application US/08471052A
CC	Patent No. 5625033
CC	GENERAL INFORMATION:
CC	APPLICANT: Kay, B. K.
CC	APPLICANT: Fowlkes, D. M.
CC	TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC	NUMBER OF SEQUENCES: 166
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Pennie & Edmonds
CC	STREET: 1155 Avenue of the Americas
CC	CITY: New York
CC	STATE: New York
CC	COUNTRY: U.S.A.
CC	ZIP: 10036-2711
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/471,052A
CC	FILING DATE: 06-JUNE-1995
CC	CLASSIFICATION: 530
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Misrock, S. Leslie
CC	REGISTRATION NUMBER: 18,872
CC	REFERENCE/DOCKET NUMBER: 1101-179
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: 212 790-9090
CC	TELEFAX: 212 869-8864/9741
CC	TELEX: 66141 PENNIE
CC	INFORMATION FOR SEQ ID NO: 142:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 69 bases
CC	TYPE: nucleic acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: unknown
CC	MOLECULE TYPE: DNA
SQ	SEQUENCE 69 BP; 2 A; 4 C; 6 G; 2 T; 55 OTHER.
Query Match      0.7%; Score 25; DB 1; Length 69;	
Best Local Similarity 11.38; Pred.No. 2.91e-01;	
Matches	7; Conservative 18; Mismatches 37; Indels 0; Gaps 0;
Db	7 AGNNBNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBCCA 66    :
Qy	1433 AGCAGCCCCCTGGCAGGTGTACGGCTTCTGCGGCCCTGCTGCGCCGCCA 1492    :
Db	67 GG 68 
Qy	1493 GG 1494

CC	STATE: New York	
CC	COUNTRY: USA	
CC	ZIP: 10036	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: PCT/US95/11934	
CC	FILING DATE: 20-SEP-1995	
CC	CLASSIFICATION:	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Misrock, S. Leslie	
CC	REGISTRATION NUMBER: 18,872	
CC	REFERENCE/DOCKET NUMBER: 1101-196-228	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: (212) 790-9090	
CC	TELEFAX: (212) 869-9741/8664	
CC	TELEX: 66141 PENNIE	
CC	INFORMATION FOR SEQ ID NO: 100:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 74 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: DNA (genomic)	
CC	SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.	
CC	Query Match	0.78; Score 25; DB 2; Length 74;
CC	Best Local Similarity	7.7%; Pred. No. 2.91e-01;
CC	Matches	5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;
Db	6 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 65	
Qy	946 CCACTCCACCACTCCGTGGCGGCAGCACCGGCGGCCCATCCACATCGCGGCC 1005	
Db	56 ACCAC 70	
Qy	1006 ACCAC 1010	
RESULT	10	
ID	PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.	
AC	xxxxxx	
DT		
DE	Sequence 100, Application PC/TUS9511934	
CC	Sequence 100, Application PC/TUS9511934	
CC	GENERAL INFORMATION:	
CC	APPLICANT: CytoGen Corporation	
CC	TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From	
CC	TITLE OF INVENTION: Peptide Libraries	
CC	NUMBER OF SEQUENCES: 103	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Pennie & Edmonds	
CC	STREET: 1155 Avenue of the Americas	
CC	CITY: New York	
CC	STATE: New York	
CC	COUNTRY: USA	
CC	ZIP: 10036	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: PCT/US95/11934	
CC	FILING DATE: 20-SEP-1995	
CC	CLASSIFICATION:	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Misrock, S. Leslie	
CC	REGISTRATION NUMBER: 18,872	
CC	REFERENCE/DOCKET NUMBER: 1101-196-228	

## CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (212) 790-9090

CC TELEFAX: (212) 869-9741/8864

CC TELEX: 66141 PENNIE

CC INFORMATION FOR SEQ ID NO: 100:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 74 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.

## SQ Query Match

SQ Best Local Similarity 0.7%; Score 26; DB 2; Length 74;

SQ Matches 7; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db 3 AGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 62

Cc 542 AGAGCGGCACGGTGCAGCAGGTGAACACGACGTCTCGCCACGCGCGCAGCA 483

Db 63 VNNACCAAC 70

Cc 482 GCCCCAC 475

## RESULT 11

ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.

AC xxxxxx

DE Sequence 94, Application PC/TUS95/11934

DE Sequence 94, Application PC/TUS95/11934

CC GENERAL INFORMATION:

CC APPLICANT: Cytogen Corporation

CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From

CC NUMBER OF SEQUENCES: 103

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pennie &amp; Edmonds

CC STREET: 1155 Avenue of the Americas

CC CITY: New York

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10036

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/11934

CC FILING DATE: 20-SEP-1995

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Mistrock, S. Leslie

CC REGISTRATION NUMBER: 18,872

CC REFERENCE/DOCKET NUMBER: 1101-196-228

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (212) 790-9090

CC TELEFAX: (212) 869-9741/8864

CC TELEX: 66141 PENNIE

CC INFORMATION FOR SEQ ID NO: 94:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 74 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

## SQ Query Match

SQ Best Local Similarity 0.7%; Score 26; DB 2; Length 74;

SQ Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Cc 3 AGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 62

Cc 542 AGAGCGGCACGGTGCAGCAGGTGAACACGACGTCTCGCCACGCGCGCAGCA 483

Db 63 VNNACCAAC 70

Cc 482 GCCCCAC 475

## RESULT 12

ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.

AC xxxxxx

DE Sequence 99, Application PC/TUS95/11934

DE Sequence 99, Application PC/TUS95/11934

CC GENERAL INFORMATION:

CC APPLICANT: Cytogen Corporation

CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From

CC NUMBER OF SEQUENCES: 103

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pennie &amp; Edmonds

CC STREET: 1155 Avenue of the Americas

CC CITY: New York

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10036

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/11934

CC FILING DATE: 20-SEP-1995

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Mistrock, S. Leslie

CC REGISTRATION NUMBER: 18,872

CC REFERENCE/DOCKET NUMBER: 1101-196-228

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (212) 790-9090

CC TELEFAX: (212) 869-9741/8864

CC TELEX: 66141 PENNIE

CC INFORMATION FOR SEQ ID NO: 99:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 75 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC SEQUENCE 75 BP; 1 A; 1 C; 7 G; 5 T; 61 OTHER.

## SQ Query Match

SQ Best Local Similarity 0.7%; Score 25; DB 2; Length 75;

SQ Matches 7; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 4 AGNN 63

Cc 168 AGGGCTGGCGGTGGTGCAGCGGGACCCGCGGCTTTCCGCGCTGGTGGCCAGT 227

Db 64 NNBGGTTGTG 73

Cc 228 GCCTGGTGTG 237

## RESULT 13

ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.

AC xxxxxx

DE Sequence 99, Application PC/TUS95/11934

DE Sequence 99, Application PC/TUS95/11934



```
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.

Query Match      0.7%; Score 25; DB 2; Length 81;
Best Local Similarity 7.7%; Pred.No. 2.91e-01;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db 13 VNNYNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 72
Qy 946 CCACTCCACCCATCCGTGGCGCCGACACACCGGGGGCCCCCATCCACATCGCGGC 1005
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 ACCAC 77
Qy 1006 ACCAC 1010
      |||||
```

Search completed: Fri Dec 25 06:26:06 1998  
Job time : 505 secs.

\*\*\*\*\*

MPERCH\_PP

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 07:42:45 1998; MasPar time 55.38 Seconds  
Tabular output not generated. 780.716 Million cell updates/sec

Title: >US-08-951-733-20  
Description: (1-1154) from US08951733.pep  
Perfect Score: 8624  
Sequence: 1 HASGQRCVLLRTWEALAPAT.....TALEAANPALPSDFKTLTD 1154

Scoring table: PAM 150  
Gap 11

Searched: 116738 segs, 37463448 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir58  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 54.465; Variance 114.185; scale 0.477

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	314	3.6	884	2	S53396 telomerase catalytic	1.90e-31
2	162	1.9	660	1	QBEB3 BHLF1 protein - human	4.72e-07
3	153	1.8	504	2	JC1306 virion protein homolo	8.51e-06
4	136	1.6	240	2	B24264 proline-rich protein	1.59e-03
5	141	1.6	276	2	B38985 hypothetical protein	2.53e-04
6	142	1.6	310	1	PIHUSD salivary proline-rich	2.61e-04
7	136	1.6	316	2	S16681 homeotic protein - hu	1.59e-03
8	137	1.6	628	2	S01955 hypothetical protein, 1.18e-03	
9	129	1.5	264	2	D34768 ORF4 protein - Orf vi	1.24e-02
10	131	1.5	300	2	S19560 proline-rich protein	6.92e-03
11	126	1.5	302	2	S11790 nodulation protein no	2.92e-02
12	129	1.5	309	2	S10889 proline-rich protein	1.24e-02
13	132	1.5	317	2	A28986 proline-rich protein	5.17e-02
14	127	1.5	566	2	S22933 testis-specific prote	2.19e-02
15	132	1.5	924	2	S27923 gene LF3 protein - hu	5.17e-03
16	131	1.5	1106	2	J00405 hypothetical 119.5K P	6.92e-03
17	130	1.5	1236	2	E70977 hypothetical protein	9.26e-03
18	117	1.4	204	2	A39086 proline-rich protein	3.53e-01
19	125	1.4	227	2	C29149 proline-rich protein	3.87e-02
20	123	1.4	240	2	A24264 proline-rich protein	6.79e-02
21	123	1.4	256	2	A50533 tumor-associated anti	6.79e-02
22	119	1.4	322	1	S00034 nodulation protein no	2.69e-01
23	119	1.4	347	2	S10571 epithelial tumor anti	2.05e-01

24	121	1.4	358	1	WMBE38 infected cell protein	1.18e-01
25	125	1.4	381	2	S16506 hypothetical protein	3.87e-02
26	125	1.4	402	2	A45056 osteogenic protein 2	3.87e-02
27	123	1.4	403	2	S52796 prpL2 protein - human	6.79e-02
28	118	1.4	431	2	S09824 hypothetical protein	2.69e-01
29	123	1.4	439	2	S51939 chitinase (EC 3.2.1.1	6.79e-02
30	124	1.4	464	2	S22697 extensin - Volvox car	5.13e-02
31	119	1.4	535	2	S10372 epithelial tumor anti	2.05e-01
32	114	1.3	202	2	B36795 hypothetical protein	7.86e-01
33	115	1.3	260	2	S22373 proline-rich protein	6.03e-01
34	116	1.3	270	2	S34361 miaE protein - Salmon	4.62e-01
35	116	1.3	295	2	B48013 proline-rich proteogl	4.62e-01
36	114	1.3	301	2	E29149 proline-rich protein	7.86e-01
37	114	1.3	321	1	A28663 nodulation protein no	7.86e-01
38	115	1.3	338	2	JC5707 HYA22 protein - human	6.03e-01
39	114	1.3	373	2	S54545 hypothetical protein	7.86e-01
40	115	1.3	392	1	PIHUB6 salivary proline-rich	6.03e-01
41	114	1.3	408	2	A56186 cyclin E - African cl	7.86e-01
42	113	1.3	704	2	A30411 synapsin Ia - rat	1.02e-00
43	114	1.3	705	2	A35363 synapsin I splice for	7.86e-01
44	113	1.3	1255	2	B35175 epistatin B - human	1.02e-00
45	115	1.3	3149	1	QBEB8 BPLF1 protein - human	6.03e-01

ALIGNMENTS

RESULT 1  
ENTRY #type complete  
TITLE telomerase catalytic chain EST2 - yeast (Saccharomyces cerevisiae)  
ALTERNATE\_NAMES protein L8543.12; protein YLR318w  
ORGANISM #formal\_name Saccharomyces cerevisiae  
DATE 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 13-Sep-1998  
ACCESSIONS S53396  
REFERENCE S53390  
#authors Du, Z.  
#submission submitted to the EMBL Data Library, February 1995  
#description The sequence of S. cerevisiae cosmid 8543.  
#accession S53396  
#molecule\_type DNA  
#residues 1-884 #label DUZ  
##cross-references EMBL:U20618; NID:92258165; PID:9662136; MIPS:YLR318w  
##experimental\_source strain S288C (AB972)

GENETICS

#gene SGD:EST2

##cross-references SGD:S0004310; MIPS:YLR318w

#map\_position 12R

SUMMARY #length 884 #molecular-weight 102662 #checksum 7604

Query Match 3.6%; Score 314; DB 2; Length 884;

Best Local Similarity 24.3%; Pred. No. 1.90e-31;

Matches 118; Conservative 122; Mismatches 202; Indels 44; Gaps 37;





```
sequence IS1222
#formal_name Enterobacter agglomerans
#3-Nov-1995 #sequence_revision 03-Nov-1995 #text_change
17-Jul-1998
ACCESSIONS
B38965; S42922
REFERENCE
#authors
Steibl, H.D.; Lewecke, F.M.
#journal
Gene (1995) 156:37-42
#title
IS1222: analysis and distribution of a new insertion sequence
in Enterobacter agglomerans 339.
#accession
B38965
#status
preliminary
#molecule_type DNA
#residues
1-276 #label STE
#cross-references GB:X78052; NID:g459246; PID:g459248
#experimental_source strain Kleeberger, 1983
GENETICS
#mobile_element insertion sequence IS1222
#length 276 #molecular_weight 31718 #checksum 3927
SUMMARY
Query Match 1.6%; Score 141; DB 2; Length 276;
Best Local Similarity 33.7%; Pred. No. 3.53e-04;
Matches 30; Conservative 23; Mismatches 30; Indels 6; Gaps 6;
Db 42 ITLALERR-FCYRR-IMOLLREGLVNHKRVYRLYHLSGLGVKRRR-RKGLATERL 98
QY 585 VTETFOKNRFLFYKRSWKLQSIGIQHLKRV-QLRSLSEAEVQHQREARPAULTSRL 643
Db 99 PLL-RPAAPNLTWSDDFVMDALATGRIK 126
QY 644 RFIKPDGLRPVIMDYVGA-RFRRREK 671
RESULT 6
ENTRY
TITLE
PIHUSD #type complete
salivary proline-rich glycoprotein precursor PRB4 (large
allele) - human
CONTAINS
basic proline-rich protein IB-5; proline-rich peptide P-D
ORGANISM
#formal_name Homo sapiens #common_name man
#length 1984 #sequence_revision 12-Apr-1996 #text_change
05-Sep-1997
ACCESSIONS
S03176; S03175; S10890; D25372; E38355; A03295; A61294;
S62891
S02127
REFERENCE
#authors
Lyons, K.M.; Stein, J.H.; Smithies, O.
#journal
Genetics (1988) 120:267-278
#title
Length polymorphisms in human proline-rich protein genes
generated by intragenic unequal crossing over.
#cross-references MUID:89121440
#accession
S03176
#status
translation not shown
#molecule_type DNA
#residues
35-310 #label LY1
#cross-references EMBL:X07715
#note
large allele
#accession
S03175
#status
translation not shown
#molecule_type DNA
#residues
35-36, 'E', '38-112, 155-310 #label LY2
#cross-references EMBL:X07704
#note
medium allele
#accession
S10890
#status
preliminary; translation not shown
#molecule_type DNA
#residues
1-38, 60-112, 'T', '114-115, 'P', '117-121, 185-271, 'A', 273-310
#label LY3
#cross-references EMBL:X07882; NID:g35647; PID:g296670
A92492
REFERENCE
#authors
Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
#journal
J. Biol. Chem. (1985) 260:11123-11130
#title
Differential RNA splicing and post-translational cleavages in
the human salivary proline-rich protein gene system.
#cross-references MUID:85289325
```

```
#accession D25372
#molecule_type mRNA
#residues
1-36, 'E', '38-112, 'T', '114-115, 'P', '117-121, 185-271, 'A',
273-310 #label MAE
REFERENCE
A38355
#authors
Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
#journal
Biochemistry (1991) 30:3351-3356
#title
Basic proline-rich proteins from human parotid saliva:
relationships of the covalent structures of ten proteins
from a single individual.
#cross-references MUID:91190884
#accession
E38355
#molecule_type protein
#residues
241-254, 'KN', 257-310 #label KAU
REFERENCE
A03295
#authors
Saitoh, E.; Isemura, S.; Sanada, K.
#journal
J. Biochem. (1983) 93:495-502
#title
Complete amino acid sequence of a basic proline-rich peptide,
P-D, from human parotid saliva.
#cross-references MUID:83186122
#accession
A03295
#molecule_type protein
#residues
241-310 #label SAI
REFERENCE
A61294
#authors
Shimomura, H.; Kanai, Y.; Sanada, K.
#journal
J. Biochem. (1983) 93:857-863
#title
Amino acid sequences of glycopeptides obtained from basic
proline-rich glycoprotein of human parotid saliva.
#accession
A61294
#molecule_type protein
#residues
54-57, 'E', 59-73, 'R', 82-101 #label SHI
REFERENCE
S62891
#authors
Charlton, A.J.; Baxter, N.J.; Lilley, T.H.; Haslam, E.;
McDonald, C.J.; Williamson, M.P.
#journal
FEBS Lett. (1996) 382:289-292
#title
Tannin interactions with a full-length human salivary
proline-rich protein display a stronger affinity than with
single proline-rich repeats.
#accession
S62891
#molecule_type protein
#residues
241-252 #label CHA
#note
amino end of peptide designated basic proline-rich
protein IB-5
#note
it is unclear from the peptide sequence whether this is
a product of the PRB2 (PIR:PIHUPF) or PRB4 (this
entry) gene
GENETICS
#gene
GDB:PRB4
#cross-references GDB:119514; OMIM:180990
#map_position 12p13.2-12p13.2
#introns
22/1; 34/1
#note
the list of introns may be incomplete
CLASSIFICATION
#superfamily proline-rich protein
KEYWORDS
glycoprotein; saliva; tandem repeat
FEATURE
1-16
241-310
66,87,171
108,150,192,213,
234
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY
#length 310 #molecular_weight 31351 #checksum 3960
Query Match 1.6%; Score 142; DB 1; Length 310;
Best Local Similarity 28.2%; Pred. No. 2.61e-04;
Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;
Db 168 OQGNOSQ-PPHPGKXPERPP-OGG-NOSHRPPPP-CKPER-PPQGNOSQ-P-PP 220
QY 199 QLGAATQARPPPHASGRRRLGCRANWNSVRAGVPLGLPAPGARRRGGSASRLPLPK 258
```



```

##status      preliminary
##molecule_type DNA
##residues    1-300  #label ROB
##cross-references GB:X58438; NID:g53181; PID:g53182
REFERENCE
S22570
  authors      Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
  #journal     Nucleic Acids Res. (1991) 19:5205-5211
  #title       The mouse proline-rich protein MP6 promoter binds
               isoprenaline-inducible parotid nuclear proteins via a
               highly conserved NFkB/rel-like site.
  #cross-references MUID:92020206
  #accession   S22570
  ##molecule_type DNA
  ##residues    1-14  #label RO2
  ##cross-references EMBL:X61126
  CLASSIFICATION #superfamily proline-rich protein
  SUMMARY        #length 300 #molecular-weight 31129 #checksum 8443

Query Match      1.5%; Score 131; DB 2; Length 300;
Best Local Similarity 25.2%; Pred. No. 6.92e-03;
Matches 34; Conservative 40; Mismatches 54; Indels 7; Gaps 7;

Db 131 PGNOGPPPOG-GPOQRPP-OPGNQCGPPPPGGPQQRPPOPGGNOGPP-OGGPHP-PPR 186
      :: ||||| :: || :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ||
QY 202 ATOARPPPHASGPPRRLLGGERAWNHVRAGVPLGLPAGCARRRGSGASRSLLPLRRPR 261
      GAAPEPRTVVGSGWAHPGRTGSPDRGFCVVSPPARPA-EAETSLEGALSGTRHSPS 320

Db 187 PGNOGPPPOGPPQQRTPQGNQCGPPPOQGG-POAPRPNGNOGPPPGQPGQPPRTGNOQ 245
      : : : : : || :: || :: || :: || :: || :: || :: || :: || :: || ::
QY 262 RGAAPERTVVGSGWAHPGRTGSPDRGFCVVSPPARPA-EAETSLEGALSGTRHSPS 320
      GAAPEPRTVVGSGWAHPGRTGSPDRGFCVVSPPARPA-EAETSLEGALSGTRHSPS 320

Db 246 -GPPPOGPGQPPRP 259
      | : : || : ||
QY 321 VGRQHHAGPPESTRP 335

RESULT 11
ENTRY   S11790 #type complete
TITLE   modulation protein nodB3 - Rhizobium leguminosarum bv.
        phaseoli
ORGANISM #formal_name Rhizobium leguminosarum bv. phaseoli
DATE     21-Nov-1993 #sequence_revision 13-Jan-1995 #text_change
        16-Feb-1997
ACCESSIONS S11790
REFERENCE   S11786
  authors    Davis, E.O.; Johnston, A.W.B.
  #journal    Mol. Microbiol. (1990) 4:921-932
  #title      Analysis of three nodD genes in Rhizobium leguminosarum
               biovar phaseoli; nodB1 is preceded by nolE, a gene whose
               product is secreted from the cytoplasm.
  #cross-references MUID:91014692
  #accession  S11790
  ##status    not compared with conceptual translation
  ##molecule_type DNA
  ##residues    1-302  #label DAV
  #gene         nodB3
  CLASSIFICATION #superfamily regulatory protein lysR
  KEYWORDS       DNA binding; transcription regulation
  SUMMARY        #length 302 #molecular-weight 34274 #checksum 3819

Query Match      1.5%; Score 126; DB 2; Length 302;
Best Local Similarity 31.4%; Pred. No. 2.92e-02;
Matches 38; Conservative 28; Mismatches 46; Indels 9; Gaps 9;

Db 17 LMERNLTAARSNLNSOPAMSAVRKL-RSYFRDEL-FTMRGREFVTP-R-AEDLAPA 72
      | : : || : : : : || || || || : : | : | : | : | : | : | : | : |
QY 9 LLRTWEALAPATP-AMPRAPRCRAVRSLLRASHYREVLPATFVRRLLGPGQWRLLVQRGDP 67
      || || || || || : : || || || || || || || || || || || || || || ||

Db 73 IREALQHLRLNIIPWDXFTPDSDRHR-YSLCDFTVTVLFLQKILERLAREAPG1SFDL 130
      || || || || || : : || || || || || || || || || || || || || || ||

QY 68 AFRALVAQCLVCVPWDA-RPPPAAPSPSRQVSKELVARVL-QRLCERGAKNVLATGFAL 125
      || || || || || : : || || || || || || || || || || || || || || ||

Db 131 L 131

```

Best Local Similarity 26.4%; Pred. No. 5.17e-03;  
Matches 39; Conservative 38; Mismatches 64; Indels 7; Gaps 6;  
Db 37 SGSPRPVNGSQGPPPGGQPRP-PQGPPPGGQPRP-PQGPPPGGQPRP-PQGP 94  
QY 202 AATQARPPPHAS--GPRRLGCE-RAWNHSVREAGVPLGLPAPGARRRGGASRSLSPLPK 258  
Db 95 PPGGQPRPQPPPGG-PQPRPQPPPGGQPRPQPPPGGQPRPQPPPGGQPRPQPPPP 153  
QY 259 RPRGAPEPRTVPVQGSWAHPGRTGRGSDRGFCVVSAPAR-PAEDATSLGALSCTRHS 317  
Db 154 GGPQPRPQPPPGGQPRPQPPPGGQPRP 181  
QY 318 HPSVGRQHAGPSTSRPRPMDTPCPP 345

RESULT 14  
ENTRY S22933 #type complete  
TITLE testis-specific protein Bsl3 - mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 10-Sep-1997  
ACCESSIONS S22933  
REFERENCE S22933  
#authors Mazarakis, N.D.; Nelki, D.; Lyon, M.F.; Ruddy, S.; Evans, E.P.; Freemont, P.; Dudley, K.  
#journal Development (1991) 111:561-571  
#title Isolation and characterisation of a testis-expressed developmentally regulated gene from the distal inversion of the mouse t-complex.  
#cross-references MUID:91372153  
#accession S22933  
#molecule\_type mRNA  
#residues 1-566 #label MAZ  
#cross-references EMBL:X52128; NID:g54852; PID:g54853  
SUMMARY #length 566 #molecular\_weight 61970 #checksum 3815

Query Match 1.5%; Score 127; DB 2; Length 566;  
Best Local Similarity 35.7%; Pred. No. 2.19e-02;  
Matches 30; Conservative 19; Mismatches 30; Indels 5; Gaps 5;  
Db 34 SARGTRVGTAVRPPSPQPRGCAVKTAPRGVGHGLRTGPTSRCPQSARAKLPS 93  
QY 239 PAPGARRRGGASRSLSPL-PKPRRGAPEPRTVPVQGS-WAHP-GRTRGSDRG-FCV 294  
Db 94 VTRGAPLPSPG-KHGLGTPSSH 116  
QY 295 VSPARAEATSLGALSCTRHS 318

RESULT 15  
ENTRY S27923 #type complete  
TITLE gene ICF3 protein - human herpesvirus 4  
ORGANISM #formal\_name human herpesvirus 4, Epstein-Barr virus  
DATE 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 08-Sep-1997  
ACCESSIONS S27923  
REFERENCE S27923  
#authors Parker, B.D.; Bankier, A.; Satchwell, S.; Barrell, B.; Farrell, P.J.  
#submission submitted to the EMBL Data Library, August 1990  
#description Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region.  
#accession S27923  
#status preliminary  
#molecule\_type DNA  
#residues 1-924 #label PAR  
#cross-references EMBL:M35547; NID:g330420; PID:g330421  
SUMMARY #length 924 #molecular\_weight 94304 #checksum 8997

Query Match 1.5%; Score 132; DB 2; Length 924;  
Best Local Similarity 28.2%; Pred. No. 5.17e-03;  
Matches 37; Conservative 33; Mismatches 52; Indels 9; Gaps 9;

Db 38 AAPRAGPEPRTLRQATPRRSGAADPADPVGHFAA-PRAPGPEPRTLRQ-PATPRRSGA 95  
QY 206 ARPPPHASGPRRLGCERAWNHSVREAGVPLGLPAPGARRRGGASRSLSPLPKRPRGAA 265  
Db 96 ADPA-DPVGHPA-A-P-RAPGPEPRTLRQATPRRSGAADPADPVGHFAA-PRAPGPEPRT 151  
QY 266 PEPRTVPVQGSWAHPGRTGRGSDRG-FCVVSAPAR-AEEATSLGALSCTRHSHP-SVG 322  
Db 152 RLQATPRRSG 162  
QY 323 ROHHAGPESTS 333

Search completed: Thu Dec 24 07:48:05 1998  
Job time : 320 secs.

\*\*\*\*\*  
M O S R L H  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Dec 24 07:48:22 1998; MasPar time 38.16 Seconds  
811.688 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-951-733-20  
Description: (1-1154) from US08951733.pep  
Perfect Score: 8624  
Sequence: 1 HASGQRCLVLTWEALPAT.....TALEAAANPALPSDFKTLTD 1154  
Scoring table: PAM 150  
Gap 11  
Searched: 74019 seqs, 26840295 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: swiss-prot35  
1:swissprot  
Statistics: Mean 55.898; Variance 96.260; scale 0.581

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	162	1.9	660	1	YHL1_EBV	4.64e-09
2	153	1.8	504	1	ATIN_HSVBP	1.47e-07
3	150	1.7	234	1	PRPM_HUMAN	4.56e-07
4	142	1.6	276	1	PRPL_HUMAN	8.72e-06
5	137	1.6	628	1	V70K_TMYC	5.26e-05
6	133	1.5	247	1	PRP4_HUMAN	2.15e-04
7	131	1.5	296	1	PRP3_MOUSE	4.31e-04
8	126	1.5	302	1	NOD3_RHLPL	2.37e-03
9	127	1.5	566	1	TS13_MOUSE	1.69e-03
10	125	1.4	261	1	PRP2_MOUSE	3.32e-03
11	121	1.4	322	1	NOD1_RHISN	1.25e-02
12	118	1.4	399	1	BM8A_MOUSE	6.23e-02
13	125	1.4	402	1	BMF8_HUMAN	3.30e-02
14	118	1.4	431	1	UL61_HCMVA	3.30e-02
15	125	1.4	501	1	MESENCHYME FORK HEAD P	3.32e-03
16	118	1.4	514	1	VE2_HPVO5	3.30e-02
17	114	1.3	202	1	VG01_HSVBP	1.17e-01
18	116	1.3	270	1	MAIE_SALTY	6.23e-02
19	115	1.3	279	1	Y091_NPVOP	8.54e-02
20	109	1.3	301	1	NDO3_RHILQ	5.40e-01
21	114	1.3	321	1	NOD1_BRAJA	1.17e-01
22	109	1.3	324	1	MATD_NEUCR	5.40e-01
23	115	1.3	331	1	PRP1_HUMAN	8.54e-02

DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE ALPHA TRANS-INDUCING PROTEIN (ALPHA-TIF).  
OS BOVINE HERPESVIRUS TYPE 1 (STRAIN P8-2).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93012995.  
RA CARPENTER D.E., MISRA V.;  
RL GENE 119:259-263(1992).  
CC -1- FUNCTION: RESPONSIBLE FOR TRANSCRIPTIONAL ACTIVATION OF IMMEDIATE  
CC -1- EARLY PROMOTERS (ALPHA GENES) (BY SIMILARITY).  
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ALPHA TRANS-INDUCING PROTEIN.  
DR EMBL; Z11610; E264419; -.  
DR PIR; S24229; S24229.  
DR PIR; JCI1306; JCI1306.  
KW TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR; DNA-BINDING.  
SQ SEQUENCE 504 AA; 54028 MW; 79F42020 CRC32;  
  
Query Match 1.8%; Score 153; DB 1; Length 504;  
Best Local Similarity 31.3%; Pred. No. 1.47e-07;  
Matches 46; Conservative 33; Mismatches 56; Indels 12; Gaps 11;  
  
Db 351 APAEAGGWRSGSTRTRGAAARSTTGLQRPCCGPRRAKC-CRATP-RQLR--ARGE 406  
QY 208 PPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSASRLPLKRRRGAPE 267  
Db 407 PRHTS-GSGAFSQ-GRPRGVRCLGWACKARSGPARGPGSPVRSGLGLSR-ARGSPGP 463  
QY 268 PERTPVGGGWAHPGTRGTPSDR-GF-CVV-S-PARAEAEATSLGALSCTRHSHPVSGR 323  
Db 464 GPACGFSRAGRRRRASPNP-FGGT 489  
QY 324 QHHAGPSTSRPPRPWDTPCPPVYAE 350

RESULT 3  
ID PRPM\_HUMAN STANDARD; PRT; 234 AA.  
AC P10161; P02813;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE M) (CONTAINS: PEPTIDE P-D)  
DE (FRAGMENT).  
GN PRB4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89121439.  
RA LYONS K.M., STEIN J.H., SMITHIES O.;  
RL GENETICS 120:255-265(1988).  
RN [2]  
RP SEQUENCE OF 165-234.  
RX MEDLINE; 83186122.  
RA SAITOH E., ISEMURA S., SANADA K.;  
RL J. BIOCHEM. 93:495-502(1983).  
DR EMBL; X07704; E265547; -.  
DR PIR; A03295; PIHUSD.  
DR PIR; S03175; S03175.  
DR HSSP; P19999; ICLG.  
DR MIM; 168730; -.  
DR MIM; 180990; -.  
KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.  
FT NON\_TER 1 1  
FT CHAIN 165 234 PEPTIDE P-D.  
SQ SEQUENCE 234 AA; 23676 MW; ED2B4ADC CRC32;  
  
Query Match 1.7%; Score 150; DB 1; Length 234;  
Best Local Similarity 29.3%; Pred. No. 4.56e-07;  
Matches 43; Conservative 41; Mismatches 53; Indels 10; Gaps 10;

Db 88 GPPP-OGGNQSQG-PPPHGKPERPPP-OGG-NQSHRPPPPP-GKPER-PPPOGGNQSOG 141  
QY 194 GPPLYQLGAATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSASRS 253  
Db 142 -P-PPHPGPEGPQEGNKSARSPPGKPGQPGQEGNKPQGGPPPGKPGQGGPPPGGN 199  
QY 254 LPLPKRRGAAPERTPVGGGWAHPGTRGTPSDR-GFCVVSPARPAE-EATSLGAL 311  
Db 200 PQQPQAPAGKPGQGGPPPGGRRPRP 226  
QY 312 SGRHSHPSVGRHAGPSTSRPRP 338  
  
RESULT 4  
ID PRPL\_HUMAN STANDARD; PRT; 276 AA.  
AC P10162; P02813;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE K) (CONTAINS: PEPTIDE P-D)  
DE (FRAGMENT).  
GN PRB4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89121439.  
RA LYONS K.M., STEIN J.H., SMITHIES O.;  
RL GENETICS 120:255-265(1988).  
RN [2]  
RP SEQUENCE OF 207-276.  
RX MEDLINE; 83186122.  
RA SAITOH E., ISEMURA S., SANADA K.;  
RL J. BIOCHEM. 93:495-502(1983).  
DR EMBL; X07715; E4806; ALT\_SEQ.  
DR PIR; A03295; PIHUSD.  
DR PIR; S03176; S03176.  
DR HSSP; P19999; ICLG.  
DR MIM; 168730; -.  
DR MIM; 180990; -.  
KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.  
FT NON\_TER 1 1  
FT CHAIN 207 276 PEPTIDE P-D.  
SQ SEQUENCE 276 AA; 27816 MW; 4838945A CRC32;  
  
Query Match 1.6%; Score 142; DB 1; Length 276;  
Best Local Similarity 28.2%; Pred. No. 8.72e-06;  
Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;  
  
Db 134 QGGNQSOG-PPPHGKPERPPP-OGG-NQSHRPPPPP-GKPER-PPPOGGNQSOG-P-PP 186  
QY 199 QLGAATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSASRSPLPK 258  
Db 187 HPCKPEGPPPGNKSARSPPGKPGQGGPPPGNKPQGGPPPGKPGQGGPPPGGNQQPQ 246  
QY 259 RPRGAAPERTPVGGGWAHPGTRGTPSDR-GFCVVSPARPAE-EATSLGALSGRH 316  
Db 247 APGAKPGQGGPPPGGRRPRP 268  
QY 317 SHPSVGRHAGPSTSRPRP 338  
  
RESULT 5  
ID V70K\_TYMV STANDARD; PRT; 628 AA.  
AC P10357;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)  
DE 69 KD PROTEIN.  
OS TURNIP YELLOW MOSAIC VIRUS.  
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; TYMOVIRIDAE.  
RN [1]

```
Db      102 PPP-OGNQSQGTPTTTPGK-PERPPP-OQG-NOSHRPPPPP-GKPER-PPPOGNQSQG-154  
       ||| :|| :|||| | :~::|| :||| :||| :||| :||| :  
Qy     195 PPLYQLGAATQAARPPPHASGP RRRLGCERAWNHSVREAGVPLGLPAPGARRRGSARSLSL 254  
  
Db     155 P-PHPCKPGSGPEOEGNKSRSARSPGKPQQOESGNPKQPQQPPPKPGQGGPPPPAGGNP 213  
       ||| :|| :|||| :||| :||| :||| :||| :||| :||| :  
Qy    255 PLPAKRPRGAAPERTPVOGGSWAHPEGRTGRGPSDR-GFCWVSFAPAEEATSLEGA LSG 313  
  
Db     214 QQODDPAGKPDGPPPPPPGGRPPRP 239  
       :   |:  :  :||| :||||  
Qy    314 TRHSHPSVGQRH-HAGPPSTSRPPR 338  
  
RESULT 7  
ID PRP3_MOUSE STANDARD; PRY; 296 AA.  
AC P05143;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DI 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
DE PROLINE-RICH PROTEIN MP-3 (FRAGMENT).  
GN PRP.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX .MEDLINE; 86059475.  
RA ANN D.K., CARLSON D.M.;  
RL J. BIOL. CHEM. 260:15863-15872(1985).  
DR ENBL; M12100; G200549; -.  
DR HSSP; PI9399; ICLG.  
KW REPEAT; SALIVA.  
FT NON_TER 1  
SQ SEQUENCE 296 AA; 29521 MW; C61DDCG6 CRC32;
```

```

Query Match          1.5%; Score 131; DB 1; Length 296;
Best Local Similarity 26.4%; Pred. No. 4.31e-04;
Matches 39; Conservative 38; Mismatches 64; Indels 7; Gaps

Db      16  SGSPRPVNGSQGPPPPGQPPRP-PQGGPPPGGQPPRPQGGPPPGGQPPRP-PQGP 73
      :::|||::|  ||  ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Qy      202 AATQAREPPPHAS--GPRRLGLCE--RAWNHSVREAGVPLGLPAGARRRGGSASRSLPLPK 258
      :::|||::|  ||  ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db      74  PPPGQPPRPQGGPPPGG-PQRPQPGPPPPGQPPRPQGGPPPGGQPPRPQGGPPPP 132
      ||  ||  ||  ||  ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Qy      259 RPRGAAPERTFVGGSWAHGRTGSDRGFCVVSPAR-PAEETSLEGALSGTRHS 317
      ||  ||  ||  ||  ||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db      133 GGQPPRPQGGPPPPGQPPRPQGGPPPP 160
      :  :  :  ::  |||  ::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy      318 HPSVGRQHHAGPPSTSRPRWDTPCP 345
      :  :  :  ::  |||  ::  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT      8
ID  NOD3_RHILP  STANDARD;  PRT;  302 AA.
AC  P23720;
DT  01-NOV-1991 (REL. 20, CREATED)
DT  01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT  01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE  NODULATION PROTEIN D III.
GN  NODD3.
OS  RHIZOBIUM LEGUMINOSARUM (BIOVAR PHASEOLI).
OC  PLASMID SYM.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC  RHIZOBIACEAE.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-8002;
RX  MEDLINE; 91014692.
RA  DAVIS E.O., JOHNSTON A.W.B.;
RL  MOL. MICROBIOL. 4:921-932(1990).
CC  -1- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCFE GENES
CC  WHICH ENCODE OTHER NODULATION PROTEINS. NODD IS ALSO A NEGATIVE
CC  REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVONOIDS AS INDUCERS.

```

[illegible]



SQ SEQUENCE 322 AA; 36460 MW; FA5238EE CRC32;

Query Match 1.4%; Score 121; DB 1; Length 322;

Best Local Similarity 30.18; Pred. No. 1.25e-02;

Matches 28; Conservative 23; Mismatches 35; Indels 7; Gaps 7;

Db 44 LRTYFRDEL-FTWNGRELVTTP-R-ABALAPAVREALLHLSIISWDPTNPAQSDRSFR 100

QY 36 LRSHYREVLPLATFVRRLGPGWRLVORGDPAAFRALVAOCLVCPWDA-RPPPAAPSFR 94

Db 101 -II-LSEFTMLFRLVVRVAREAPAVSFELL 131

QY 95 QVSCLELVARV-LQRLCERGAKNVLAFFGALL 126

RESULT 12

ID BNP8\_MOUSE STANDARD; PRT; 399 AA.

AC P34821;

DT 01-FEB-1994 (REL. 28, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE BONE MORPHOGENETIC PROTEIN 8A PRECURSOR (BMP-8A) (OSTEOGENIC PROTEIN

2) (OP-2).

GN BMP8A OR BMP8.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; RODENTIA.

CC [1]

RN SEQUENCE FROM N.A.

RC TISSUE=EMBRYO;

RE MEDLINE; 93094231.

RA OZKAYNAK E., SCHNEGELSBERG P.N.J., JIN D.F., CLIFFORD G.M.,

RL J. BIOL. CHEM. 267:25220-25227(1992).

CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE

CC OSTEOINDUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF

CC EPITHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION

CC AND BONE HOMEOSTASIS (BY SIMILARITY).

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.

CC -1- DEVELOPMENTAL STAGE: EXTENSIVE EXPRESSION FOUND IN 8-DAY EMBRYOS,

CC FELL DRASTICALLY IN 10-DAY EMBRYOS AND VIRTUALLY ABSENT IN 17-

CC DAY EMBRYOS.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; M97017; G200141; --

DR MGI; M97017; G200141; --

DR PROSITE; PS00250; TGF\_BETA; 1.

KW SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 260 POTENTIAL.

FT CHAIN 261 399 BONE MORPHOGENETIC PROTEIN 8A.

FT DISULFID 298 364 BY SIMILARITY.

FT DISULFID 327 396 BY SIMILARITY.

FT DISULFID 331 398 BY SIMILARITY.

FT DISULFID 363 363 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 155 155 POTENTIAL.

FT CARBOHYD 340 340 POTENTIAL.

SQ SEQUENCE 399 AA; 44764 MW; 7075ALED CRC32;

Query Match 1.4%; Score 118; DB 1; Length 399;

Best Local Similarity 30.78; Pred. No. 3.30e-02;

Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GLALCALGGGHPRP-RTC-PQRLG-ARERDMQREILAVLGLPGR-PPRPAQAAAR 67

QY 194 GPPLYQLGAATQAPPPHSGRRRLGCRANNSHVSREAGVPLGLPAGARRRGSASRS 253

Db 68 QP 69

QY 254 LP 255

RESULT 13

ID BMP8\_HUMAN STANDARD; PRT; 402 AA.

AC P34821;

DT 01-FEB-1994 (REL. 28, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE BONE MORPHOGENETIC PROTEIN 8A PRECURSOR (BMP-8A) (OSTEOGENIC PROTEIN 2)

DE (OP-2).

GN BMP8.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

CC [1]

RN SEQUENCE FROM N.A.

RC TISSUE=HIPPOCAMPUS;

RE MEDLINE; 93094231.

RA OZKAYNAK E., SCHNEGELSBERG P.N.J., JIN D.F., CLIFFORD G.M.,

RL J. BIOL. CHEM. 267:25220-25227(1992).

CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE

CC OSTEOINDUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF

CC EPITHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION

CC AND BONE HOMEOSTASIS (BY SIMILARITY).

CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; M97016; G189390; --

DR PIR; A45056; A45056.

DR HSSP; P08112; ITPG.

DR MIM; 602284; --

DR PROSITE; PS00250; TGF\_BETA; 1.

KW SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 263 POTENTIAL.

FT CHAIN 264 402 BONE MORPHOGENETIC PROTEIN 8.

FT DISULFID 301 367 BY SIMILARITY.

FT DISULFID 330 399 BY SIMILARITY.

FT DISULFID 334 401 BY SIMILARITY.

FT DISULFID 366 366 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 158 158 POTENTIAL.

FT CARBOHYD 343 343 POTENTIAL.

SQ SEQUENCE 402 AA; 44764 MW; 8F20C81B CRC32;

Query Match 1.4%; Score 125; DB 1; Length 402;

Best Local Similarity 39.78; Pred. No. 3.32e-03;

Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;

Db 12 GLALCALGGGHPRP-TC-PQRLG-ARERDVQREILAVLGLPGR-PPRPAAPAS 67

QY 194 GPPLYQLGAATQAR-PPPHASGPRRLGCRANNSHVSREAGVPLGLPAGARRRGSASR 252

Db 68 RLP 70

QY 253 SLP 255

RESULT 14

ID UL61\_HCMVA STANDARD; PRT; 431 AA.

AC P16818;

DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN UL61.

GN HUMAN CYTOMEGALOVIRUS (STRAIN AD169).

OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.

CC [1]

RN SEQUENCE FROM N.A.

RC MEDLINE; 90269039.

RA CHEE M.S., BANKIER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R.,

RA HORSNELL T., HUTCHISON C.A. III, KOUZARIDES T., MARTIGNETTI J.A.,

RA PREDDIE E., SATCHELL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.,

RL CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990).

DR EMBL; X17403; E27460; --

DR PIR; S09824; S09824.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 431 AA; 44309 MW; 232AB9D7 CRC32;  
Query Match 1.4%; Score 118; DB 1; Length 431;  
Best Local Similarity 32.1%; Pred. No. 3.30e-02;  
Matches 27; Conservative 20; Mismatches 31; Indels 6; Gaps 6;  
Db 268 RGGGKPLGSPRA-TDGNRDPG-A-GVPARPGREWGSSGGGTGPGGPERAAPGARP 324  
QY 211 HASGPRRLGGERAWNHSVREAGVPLGPGAGARRRGSAS-RSLPLKPRRRGAAPPE 269  
DE FKHL14)  
Db 325 TAPDGPAGRWGDPADGAPGLGRG 348  
QY 270 RTPVG-QGSWAHPGRTGFS-DRG 291

RESULT 15  
ID MFH1\_HUMAN STANDARD; PRT; 501 AA.  
AC Q99958;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MESENCHYME FORK HEAD PROTEIN 1 (MFH-1 PROTEIN) (TRANSCRIPTION FACTOR  
DE FKHL14)  
GN FKHL14 OR MFH1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97312712.  
RA MIURA N., IIDA K., KAKINUMA H., YANG X.-L., SUGIYAMA T.;  
RL GENOMICS 41:489-492(1997).  
CC -!- FUNCTION: MIGHT BE INVOLVED IN THE FORMATION OF SPECIAL  
CC MESENCHYMAL TISSUES (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.  
DR EMBL; Y08223; E303016; -.  
DR MIM; 602402; -.  
DR PROSITE; PS00657; FORK\_HEAD\_1; 1.  
DR PROSITE; PS00658; FORK\_HEAD\_2; 1.  
DR PROSITE; PS50039; FORK\_HEAD\_3; 1.  
KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.  
FT DNA\_BIND 71 162 FORK-HEAD.  
FT DOMAIN 163 167 POLY-ARG.  
FT DOMAIN 387 396 HIS-RICH.  
FT DOMAIN 397 421 ALA/PRO-RICH.  
FT DOMAIN 400 408 POLY-PRO.  
FT DOMAIN 416 422 POLY-ALA.  
SQ SEQUENCE 501 AA; 53719 MW; 99D32EFA CRC32;

Query Match 1.4%; Score 125; DB 1; Length 501;  
Best Local Similarity 28.4%; Pred. No. 3.32e-03;  
Matches 27; Conservative 28; Mismatches 38; Indels 2; Gaps 2;  
Db 367 SPLSALNLAAQEGALATGHHQHGHHPQAPPPAPQPPQPTPGGAAAAQASWYL 426  
QY 296 SPAPAEATLEGALSGTRHSPSVGRQH-HAGPPTSRRPPRPWDTPCPVYAEIKHFL 354  
DE 427 NMSGDLNHL-PGHTFAAOQTFFPNVREMFNSHRLG 460  
QY 355 YSSGDKQLRPSFLLSLRPSLTGARRLIVETIFLG 389

Search completed: Thu Dec 24 07:51:35 1998  
Job time : 193 secs.

\*\*\*\*\*  
MPSRLH (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Dec 24 07:51:54 1998; MasPar time 76.57 Seconds  
Tabular output not generated. 750.503 Million cell updates/sec

Title: >US-08-951-733-20  
Description: (1-1154) from US08951733.pep  
Perfect score: 8624  
Sequence: 1 HASGQRCVLLRTWEALAPAT.....TALEAAANPALPSDFKTLTD 1154

Scoring table: PAM 150  
Gap 11  
Searched: 165420 seqs, 49795644 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: spiremb6  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus  
Statistics: Mean 54.168; Variance 101.474; scale 0.534

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length DB ID	Description	Pred. No.		
1	8465	98.2	1132 4	O14746 TELOMERASE REVERSE TRA	0.00e+00		
2	8459	98.1	1132 4	O14783 TELOMERASE CATALYTIC S	0.00e+00		
3	5161	59.8	1122 11	O70372 TELOMERASE REVERSE TRA	0.00e+00		
4	678	7.9	988 3	O13339 TELOMERASE REVERSE TRA	1.37e-110		
5	681	7.9	989 3	O13338 TELOMERASE REVERSE TRA	3.12e-111		
6	439	5.1	1031 5	O00939 TELOMERASE SUBUNIT P12	1.97e-60		
7	424	4.9	67 11	O35432 TELOMERASE CATALYTIC S	2.25e-57		
8	314	3.6	884 3	O06163 CHROMOSOME XII COSMID	1.71e-35		
9	147	1.7	296 14	O69118 HYPOTHETICAL PROTEIN (	7.18e-06		
10	143	1.7	574 3	O36027 WISKOTT-ALDRICH SYNDRO	2.85e-05		
11	142	1.6	234 4	O00600 PAROTID 'O' PROTEIN (F	4.02e-05		
12	141	1.6	276 2	O04612 IS 1222 GENE ORF-A AND	5.65e-05		
13	136	1.6	316 4	O99076 HOMEOBOX PROTEIN (FRAG	3.03e-04		
14	140	1.6	373 2	O54226 POLYKETIDE SYNTHASE (F	7.93e-05		
15	135	1.6	585 14	O41935 HYPOTHETICAL 60.2 KD P	4.23e-04		
16	136	1.6	897 11	O70495 PLENTY-OF-PROLINES-101	3.03e-04		
17	129	1.5	264 14	O85301 HOMOLOGUE OF RETROVIRA	3.01e-03		
18	131	1.5	300 11	O61888 PROLINE RICH PROTEIN.	1.57e-03		
19	129	1.5	309 4	O04188 SALIVARY PROLINE-RICH	3.01e-03		
20	132	1.5	317 11	O62103 PROLINE RICH PROTEIN P	1.14e-03		

21	128	1.5	527 2	O53913 BETA-KETOACYL SYNTHASE	4.15e-03		
22	133	1.5	539 5	O02123 SIMILARITY TO COLLAGEN	8.18e-04		
23	132	1.5	924 14	O99307 LF3 PROTEIN.	1.14e-03		
24	130	1.5	1236 2	O06264 HYPOTHETICAL 131.9 KD	2.18e-03		
25	126	1.5	3247 14	O65553 UL36.	7.85e-03		
26	117	1.4	197 5	O17626 C04G2.8.	1.27e-01		
27	118	1.4	203 5	O91497 CODED FOR BY C. ELEGAN	9.42e-02		
28	125	1.4	227 11	O62107 PROLINE-RICH SALIVARY	1.08e-02		
29	118	1.4	236 14	O85028 SUBSTRATE OF THE PROTE	9.42e-02		
30	118	1.4	322 2	OQ53180 MPK3030 NODD1 GENE IN	9.42e-02		
31	118	1.4	405 2	O68033 EXONUCLEASE SBCD HOMOL	9.42e-02		
32	118	1.4	437 2	O52256 FLHF.	9.42e-02		
33	123	1.4	439 10	O42421 CHITINASE PRECURSOR.	2.02e-02		
34	124	1.4	464 10	O41645 EXTENSIN (FRAGMENT).	1.48e-02		
35	123	1.4	494 4	O15220 PRPL-2 PROTEIN.	2.02e-02		
36	125	1.4	503 4	O43516 WASP INTERACTING PROTE	1.08e-02		
37	118	1.4	526 14	O85027 VIRAL PROTEINASE.	9.42e-02		
38	122	1.4	657 10	O22835 SP16 ISOLOG.	2.76e-02		
39	124	1.4	820 4	O60585 SER/ARG-RELATED NUCLEA	1.48e-02		
40	115	1.3	106 14	O41981 HYPOTHETICAL 9.9 KD PR	2.32e-01		
41	115	1.3	260 11	O64306 PROLINE-RICH PROTEIN.	2.32e-01		
42	116	1.3	295 11	O07611 PROLINE-RICH PROTEOGLY	1.72e-01		
43	115	1.3	297 4	O16038 PRB1M PROTEIN (FRAGEN	2.32e-01		
44	115	1.3	382 4	O00599 CON1.	2.32e-01		
45	116	1.3	488 10	O04528 SEQUENCE OF BAC F20P5	1.72e-01		

ALIGNMENTS

RESULT	1													
ID	O14746	PRELIMINARY;	PRT;	1132	AA.									
AC	O14746;													
DT	01-JAN-1998	(TREMBUREL. 05, CREATED)												
DT	01-JAN-1998	(TREMBUREL. 05, LAST SEQUENCE UPDATE)												
DT	01-JUN-1998	(TREMBUREL. 06, LAST ANNOTATION UPDATE)												
DE	TELOMERASE REVERSE TRANSCRIPTASE.													
GN	HTRT.													
OS	HOMO SAPIENS (HUMAN).													
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;													
OC	EUTHERIA; PRIMATES.													
RN	[1]													
RP	SEQUENCE FROM N.A.													
RC	TISSUE-KIDNEY;													
RX	MEDLINE; 97400623.													
RA	NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,													
RA	LINGNER J., HARLEY C.B., CECCH T.R.;													
RL	SCIENCE 277:955-959(1997).													
DR	EMBL; AF015950; G2330017.													
KW	RNA-DIRECTED DNA POLYMERASE.													
SQ	SEQUENCE 1132 AA; 126995 MW; 2DFBEDF3 CRC32;													

Query Match				98.2%; Score 8465; DB 4; Length 1132;			
Best Local Similarity 100.0%; Pred. No. 0.00e+00;							
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Db	1	MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLRGDPAAFRALVAQCLVCPW	60				
QY	23	MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLRGDPAAFRALVAQCLVCPW	82				
Db	61	DARPPPAASFQVSCLEKELVAVLRQLRGERGAKNVLFATGFDLLDARGGPPPAFTTSVR	120				
QY	83	DARPPPAASFQVSCLEKELVAVLRQLRGERGAKNVLFATGFDLLDARGGPPPAFTTSVR	142				
Db	121	SYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLVOLGA	180				
QY	143	SYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLVOLGA	202				
Db	181	ATQARPPPHASGPRRLGGERAWNHSVRAGVPLGLPAPGARRRGSGASRSLPKRPRR	240				
QY	203	ATQARPPPHASGPRRLGGERAWNHSVRAGVPLGLPAPGARRRGSGASRSLPKRPRR	262				
Db	241	GAAPERTVVGSGWAHPGRTGSPDRGFCVVVSPARPAEATSLGALSGTRHSPSYG	300				

QY 263 GAAPERTPVGOGSWAHGCRTRGSDRGFCVVSPPARPAEATSLGALSCTRHSHPVSG 322  
Db 301 ROHAGPPSTRPRPDWTPCPVPYAEATKHFYSSGDKQLRFSFLSSLRPSLTGARRL 360  
QY 323 ROHAGPPSTRPRPDWTPCPVPYAEATKHFYSSGDKQLRFSFLSSLRPSLTGARRL 382  
Db 361 VETIFGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGNHACPYGVLLKTHCPPLRAAVT 420  
QY 383 VETIFGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGNHACPYGVLLKTHCPPLRAAVT 442  
Db 421 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 480  
QY 443 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 502  
Db 481 RHNERFLRNTKFIISLGKHAHSLQELTWKMSVRDCAWLRSPSGVGCVPAAEHLREEI 540  
QY 503 RHNERFLRNTKFIISLGKHAHSLQELTWKMSVRDCAWLRSPSGVGCVPAAEHLREEI 562  
Db 541 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRFFYRKSVWSKLSQIGIRHKLKVOLRE 600  
QY 563 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRFFYRKSVWSKLSQIGIRHKLKVOLRE 622  
Db 601 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660  
QY 623 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 682  
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 720  
QY 683 LFSVLNTERARRPGLLGASVLGLDDIHRWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 742  
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780  
QY 743 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802  
Db 781 QETSPLDADVIEQSSINAEASGLDFVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 840  
QY 803 QETSPLDADVIEQSSINAEASGLDFVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 862  
Db 841 LCSLCYGDMEKFLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRVLVRGPEYGCYVNL 900  
QY 863 LCSLCYGDMEKFLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRVLVRGPEYGCYVNL 922  
Db 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTLRTLEVDQSDYSYARTSTRASLTF 960  
QY 923 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTLRTLEVDQSDYSYARTSTRASLTF 982  
Db 961 NRGEKAGNRMRKFLGVLRKCHSLFDLDQVNSLOTVCTNIYKILLQAVRFHACVLQLP 1020  
QY 983 NRGEKAGNRMRKFLGVLRKCHSLFDLDQVNSLOTVCTNIYKILLQAVRFHACVLQLP 1042  
Db 1021 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1080  
QY 1043 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL 1102  
Db 1081 KLTHRVTYVPLGSLRTAQTLRSLKPLGTLTALAAANPALSPDFKTIID 1132  
QY 1103 KLTHRVTYVPLGSLRTAQTLRSLKPLGTLTALAAANPALSPDFKTIID 1154

RESULT 2  
ID O14783 PRELIMINARY; PRT: 1132 AA.  
AC O14783;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE TELOMERASE CATALYTIC SUBUNIT.  
GN HEST2  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97433088.

RA MEYERSON M., COUNTER C.M., EATON E.N., ELLISEN L.W., STEINER P.,  
RA CADDLE S.D., ZIAUGRA L., BEIJERSBERG R.L., DAVIDOFF M.J., LIU Q.,  
RA BACCHETTI S., HABER D.A., WEINBERG R.A.;  
RL CELL 90:785-795(1997).  
DR EMBL; AF018167; G2347129; -;  
SQ SEQUENCE 1132 AA; 126937 MW; C1E5E2AF CRC32;  
Query Match 98.1%; Score 8459; DB 4; Length 1132;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 1 MPAPPCRAVRSLLRSYREVLPATFVRLRPGQCHRLVQRGDPAAFRALVAOCLVCPW 60  
QY 23 MPAPPCRAVRSLLRSYREVLPATFVRLRPGQCHRLVQRGDPAAFRALVAOCLVCPW 82  
Db 61 DARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVLAFCFALLDARGGPPPEAFVTSVR 120  
QY 83 DARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVLAFCFALLDARGGPPPEAFVTSVR 142  
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYLQGA 180  
QY 143 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYLQGA 202  
Db 181 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARBRGGSASRSLPKRPRR 240  
QY 203 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARBRGGSASRSLPKRPRR 262  
Db 241 GAAPERTPVGOGSWAHGCRTRGSDRGFCVVSPPARPAEATSLGALSCTRHSHPVSG 300  
QY 263 GAAPERTPVGOGSWAHGCRTRGSDRGFCVVSPPARPAEATSLGALSCTRHSHPVSG 322  
Db 301 ROHAGPPSTRPRPDWTPCPVPYAEATKHFYSSGDKQLRFSFLSSLRPSLTGARRL 360  
QY 323 ROHAGPPSTRPRPDWTPCPVPYAEATKHFYSSGDKQLRFSFLSSLRPSLTGARRL 382  
Db 361 VETIFGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGNHACPYGVLLKTHCPPLRAAVT 420  
QY 383 VETIFGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGNHACPYGVLLKTHCPPLRAAVT 442  
Db 421 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 480  
QY 443 PAAGVCAREKPOGSVAAPAEEDTDPRLLVOLLRHSSPMQVYGFVRACLRRLVPPGLWGS 502  
Db 481 RHNERFLRNTKFIISLGKHAHSLQELTWKMSVRDCAWLRSPSGVGCVPAAEHLREEI 540  
QY 503 RHNERFLRNTKFIISLGKHAHSLQELTWKMSVRDCAWLRSPSGVGCVPAAEHLREEI 562  
Db 541 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRFFYRKSVWSKLSQIGIRHKLKVOLRE 600  
QY 563 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRFFYRKSVWSKLSQIGIRHKLKVOLRE 622  
Db 601 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660  
QY 623 LSEAEVQHRREARPAALLTSRLRFTPKDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 682  
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 720  
QY 683 LFSVLNTERARRPGLLGASVLGLDDIHRWRTFVLVRAQDPPPELYFVKVDVTGAYDTI 742  
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780  
QY 743 PODRLTEVIASIIKPONTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802  
Db 781 QETSPLDADVIEQSSINAEASGLDFVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 840  
QY 803 QETSPLDADVIEQSSINAEASGLDFVFLRFMCHHVRIRGKSVYOCQIGPOGSIILSTL 862  
Db 841 LCSLCYGDMEKFLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRVLVRGPEYGCYVNL 900  
QY 863 LCSLCYGDMEKFLFAGIRRDGLLLRLVDDFLVTPHLLTHAKTFLRVLVRGPEYGCYVNL 922  
Db 901 RKTVMNPFVEDEALGGTAFVQMPAHGLFPWCGLLDTLRTLEVDQSDYSYARTSTRASLTF 960

QY 923 RKTVPNFEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEQSDYSYARTSIRASLTF 982  
Db 961 NRGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1020  
QY 983 NRGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1042  
Db 1021 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGNSLAKGAAGPLPSEAWOWLCHQAFLL 1080  
QY 1043 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGNSLAKGAAGPLPSEAWOWLCHQAFLL 1102  
Db 1081 KLTHRYTYVPLGLSLTAQTLQSKRLPGTTLTALEAAANPALPSDFKTLID 1132  
QY 1103 KLTHRYTYVPLGLSLTAQTLQSKRLPGTTLTALEAAANPALPSDFKTLID 1154  
RESULT 3  
ID 070372 PRELIMINARY; PRT; 1122 AA.  
AC 070372;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DE TELOMERASE REVERSE TRANSCRIPTASE.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RP [1]  
RP SEQUENCE FROM N.A.  
RA GREENBERG R.A., ALLSOPP R.C., CHIN L., MORIN G.B., DEPINHO R.A.;  
RL ONCOGENE 0:0-0(1998).  
DR ENBL: AF051911; G3005592; -  
KW RNA-DIRECTED DNA POLYMERASE.  
SQ SEQUENCE 1122 AA; 127977 MW; 222075d6 CRC32;  
Query Match 59.8%; Score 5161; DB 11; Length 1122;  
Best Local Similarity 62.7%; Pred. No. 0.00e+00;  
Matches 718; Conservative 198; Mismatches 192; Indels 38; Gaps 24;  
Db 1 MTRAPRCRAVRSLLRSYREVWPLATFVRRLGPEGRRLVQDPGPKIYRTLVAQCLVCMHW 60  
QY 23 MPRAPRCRAVRSLLRSYREVWPLATFVRRLGPEGRRLVQDPGPKIYRTLVAQCLVCMHW 82  
Db 61 GSOPPPADLSFHOVSLKELVARVVORLCERNERNLAFGLFELLNEARGPPMAFTSVR 120  
QY 83 DARPPAPRSPFQVSKLKLVARVQLRCERGAKNVLAFGFALLDARGGPEAFTSVR 142  
Db 121 SYLPTVETLRVSGAMWLLSRVGDLLVYLLAHCALYLLVPPSCAYQVCGSPLYOICA 180  
QY 143 SYLPTVTDALRGSGANGLLRVGDVLLHLLARCALFVLVAPSCAYQVCGSPLYOIGA 202  
Db 181 TTDIWPVSASRYTRPVGRNFTWRLFLQKSSROEAPKPLALPSRGTKRHLSLST 240  
QY 203 ATQARPPPHAS-GPRRLGCE-RA-W--NH-SVR-EAGVPLGLPAPGARRRGGSASRS 253  
Db 241 VPSAKKARCYPVPRVEGPHQVLTPTSGKSWPSPARS-PEV-PT-AEKDLSS-KKVS 296  
QY 254 LPLKPRRGAAPERTPTVQGSWAHPGRTGRPS-DRGFCWSPAPAEATLEALS 312  
Db 297 DLSLS-GSVCKKXPSSTLSLPPRQNAFLRP-FITRHFYLSRGDQOERLNPSELLSN 354  
QY 313 GTRHSHFSVGRHAGPSPSTRPPRWDTPCPVYATKHFYSSGD-KEOLRFSFLLSS 371  
Db 355 LOPNLGARRLVEIFLGRSPRTSGPLCRTHLSRRYWMQRLPFOQLLVNHAECQYVRL 414  
QY 372 LRPSLTGARRLVETIFLGRSPMGTGPRRLPRLPQRYWMQRLPFLLELGNHACQYVGL 431  
Db 415 RSHCRFR---T--AN---Q--Q--VT--DALNTSPHLLMDLLKLSHPQVYGLRACL 459  
QY 432 KTHCPLRAVTPAAGVCAREKPOGSAPEEEDTPRLVQLLRHQHSPWQVYGFVRACL 491  
Db 460 CKVYSASLWGRHNRHREFKLLKFIISLQYKGLSLQELMWMKVEDCHWLRSSPGDRV 519  
QY 492 RRLVPPGLWGRHNRHREFKLLKFIISLQYKGLSLQELMWMKVEDCHWLRSSPGVGV 551

Db 520 PAEHLRERILATFLFWLMDTVVOLLRSFFITESTFQKNLFFYRKSWSKLSIGV 579  
QY 552 PAEHLRERILATFLFWLMDTVVOLLRSFFITESTFQKNLFFYRKSWSKLSIGV 611  
Db 580 ROHLERVLRELSEOEVRHODTWLAMPICRLRFIPKPNGLRPIVNMYSYSGTRALGRK 639  
QY 612 ROHLERVLRELSEOEVRHODTWLAMPICRLRFIPKPNGLRPIVNMYSYSGTRALGRK 671  
Db 640 QAQHTFORLKTLSMNLNERTKPHLGMSSVLGMNDIYRTWRAFLVRLALDQTPRMFY 699  
QY 672 RAERLTSRVKALFVLNYSERARRPGLLGASVLGDDHRAWRFTVLRVRAQDPPPELYFV 731  
Db 700 KADVTGAYDAIPOGKLVVANNIRHSESTYCIQRYAVVRDQSGQVHKHFRQVTLSD 759  
QY 732 KADVTGAYDIPQDRLTEVIASIK-PONTYCVRRYAVQKAAHGHVYKAFKSHVSTLD 790  
Db 760 LQPMGQFLKHLQSDASALRNVSIVTEQISMNESSSLSLDFLHLFLRHSVVKIGRCYT 819  
QY 791 LQPMGQFVAHQET--SPLRDVAVIEQSSLSNEASGLDFELRMCHHAVIRKSYV 848  
Db 820 QCQIGPQSSSLTLLCSLCFGDMENKLFARVQDGLLRLFRVDDFLVTPHLDQAKTFLST 879  
QY 849 QCQIGPQSSSLTLLCSLCFGDMENKLFARVQDGLLRLFRVDDFLVTPHLDQAKTFLST 908  
Db 880 LVHGVPEYGCMLNLOKTVVNFVPGTLGGAAPYQIPAHCLFPWCGLLLDTQTLVEFCDY 939  
QY 909 LVHGVPEYGCMLNLOKTVVNFVPGTLGGAAPYQIPAHCLFPWCGLLLDTQTLVEFCDY 968  
Db 940 SGVAOTSIKTSLTFSQSVFKAGTKMRNKLKLSVLKCHGLFLDLQVNSLQTVCTNIYKIFL 999  
QY 969 SSYARTSIRASLTFRNGFKAGNNRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILL 1028  
Db 1000 LQAYRFHACVQLPFPQVRKNTLFFLGIISQASOACYALVKVKNPGMTL--K-ASGSFP 1056  
QY 1029 LQAYRFHACVQLPFPQVRKNTLFFLGIISQASOACYALVKVKNPGMTL--K-ASGSFP 1088  
Db 1057 PERAHWCYQAFLLKLAHVSIVYKLLGLPLRTAQKLCRLKLPKPEATMTILKAAADPALSTD 1116  
QY 1089 SEAWMLCHQAFLLKLTNRHRTVYVPLGSLRTAQTLQSLRTPGLTTLTALEAAANPALPSD 1148  
Db 1117 FOTILD 1122  
QY 1149 FKTILD 1154  
RESULT 4  
ID 013339 PRELIMINARY; PRT; 988 AA.  
AC 013339;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DE TELOMERASE REVERSE TRANSCRIPTASE 1.  
GN TET1  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RX MEDLINE; 97400623.  
RA NAKAMURA T.N., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,  
RA LINGNER J., HARLEY C.B., CECI T.R.;  
RL SCIENCE 277:955-959(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;  
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR ENBL: AF015783; G2340168; -  
DR ENBL: AL022299; E1285360; -  
DR PFAM: PF00078; ITC.  
KW RNA-DIRECTED DNA POLYMERASE.  
SQ SEQUENCE 988 AA; 116328 MW; 39C385A7 CRC32;

Query Match 7.9%; Score 678; DB 3; Length 988;

Best Local Similarity 26.4%; Pred. No. 1.37e-110;

Matches 129; Conservative 133; Mismatches 208; Indels 18; Gaps 15;

Db 340 PNQVFAELRSLRVFPKLINGNORIFEIILKDLFTLKLRYESFSLHYLMNSIKISEI 399

QY 480 PQWYGVFVACLRRLVPPGLGWSRHRNRRFLNTKFKISLGKHAQLSLOELTWKMSVRDC 539

Db 400 EWLVLGRSNKAMCLSDPEKR-KO-IPAEFYIWLNSFIIPLOSFFYITESSDLNRRTV 457

QY 540 AWL---RRSPGVCVPAAEHLREILAKFLHLMVSVVVELLSRFFVYVTTTFOKNRLF 596

Db 458 YFRKDIW-KLRCRPFITSMKMEAFKINENNVMDTQ-KTTLPPAVIRLLPKKNTFRILT 515

QY 597 FYRKSVMKLSIGIRQLHKLVRQLRELSEAEVRQHRARPALLTSRLRFPKPDGLRPV 656

Db 516 NLKRFLIKMGSNKMVSTNQTLPVASILKHLINNESSGIPFN-LEVYMKLLTFKKDL 574

QY 657 NMDYVVGARTFRERKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRARTFV 716

Db 575 LKHRMGRK-K-YFVRIDIKSCYDRIKODLMFRIVKVKLKDPE-FVIRKYATIH-ATS 630

QY 717 LRVAQDPPPELYFVKVDVTGAYDTIPODRLTEVIASIIKPNYCVRRYAVVQKAAHG 776

Db 631 ATKNFVSEAFSYFDMVPEK--VVQLLSMKT-SDTLFDVFDYDWTKSSSEIFKMLKEHLS 687

QY 777 VRKAFKSHVSTLTDLPYMRQFVAHLQETSPLRDAVVEQSSSLNEASSGLDFVFLRFMC 836

Db 688 GHIVKIGNSQYLVKQVIGPOGSILSFCHFYMEDLIDELYSFTKKGSVLLRVVDDFLFI 747

QY 837 HHAVIRKSVQCOGIPQGSILSTLSCYGDENKLFAGIRRDG-LLRLVDDFLV 895

Db 748 TVNKKDAKFLNLSRGFEKHNFSTSLKTYVINFENSGIINNTFFNESKKR-M-PFFGF 805

QY 896 TPLHTHAKTFLRVLRGVPEYGCVVNLRKTVVNFPEDEALGGTAFVQMPAHGLFPWGL 955

Db 806 SVNMRSLD 813

QY 956 LLDTRTLE 963

RESULT 5

ID 013338 PRELIMINARY; PRT; 989 AA.

AC 013338

DT 01-JAN-1998 (TREMUREL. 05, CREATED)

DT 01-JAN-1998 (TREMUREL. 05, LAST SEQUENCE UPDATE)

DT 01-JUN-1998 (TREMUREL. 06, LAST ANNOTATION UPDATE)

DE TELOMERASE REVERSE TRANSCRIPTASE 1.

GN TR1.

OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-972H-

RX MEDLINE; 97400623.

RA NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,

RA LINGNER J., HARLEY C.B., CECCH T.R.;

RL SCIENCE 277:955-959 (1997).

DR EMBL; AF015783; G2340169; -.

DR PFAM; PF000078; rvt.

KW RNA-DIRECTED DNA POLYMERASE.

QY SEQUENCE 989 AA; FDE74202 CRC32;

Query Match 7.9%; Score 681; DB 3; Length 989;

Best Local Similarity 26.3%; Pred. No. 3.12e-111;

Matches 129; Conservative 136; Mismatches 204; Indels 21; Gaps 18;

Db 340 PNQVFAELRSLRVFPKLINGNORIFEIILKDLFTLKLRYESFSLHYLMNSIKISEI 399

QY 480 PQWYGVFVACLRRLVPPGLGWSRHRNRRFLNTKFKISLGKHAQLSLOELTWKMSVRDC 539

Db 400 EWLVLGRSNKAMCLSDPEKR-KO-IPAEFYIWLNSFIIPLOSFFYITESSDLNRRTV 457

QY 540 AWL---RRSPGVCVPAAEHLREILAKFLHLMVSVVVELLSRFFVYVTTTFOKNRLF 596

Db 458 YFRKDIW-KLRCRPFITSMKMEAFKINENNVMDTQ-KTTLPPAVIRLLPKKNTFRILT 515

QY 597 FYRKSVMKLSIGIRQLHKLVRQLRELSEAEVRQHRARPALLTSRLRFPKPDGLRPV 656

Db 516 NLKRFLIKMGSNKMVSTNQTLPVASILKHLINNESSGIPFN-LEVYMKLLTFKKDL 574

QY 657 NMDYVVGARTFRERKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRARTFV 716

Db 575 LKHRMGRK-K-YFVRIDIKSCYDRIKODLMFRIVKVKLKDPE-FVIRKYATIH-ATS 630

QY 717 LRVAQDPPPELYFVKVDVTGAYDTIPODRLTEVIASIIKPNYCVRRYAVVQKAAHG 776

Db 631 ATKNFVSEAFSYFDMVPEK--VVQLLSMKT-SDTLFDVFDYDWTKSSSEIFKMLKEHLS 687

QY 777 VRKAFKSHVSTLTDLPYMRQFVAHLQETSPLRDAVVEQSSSLNEASSGLDFVFLRFMC 836

Db 688 GHIVKIGNSQYLVKQVIGPOGSILSFCHFYMEDLIDELYSFTKKGSVLLRVVDDFLFI 747

QY 837 HHAVIRKSVQCOGIPQGSILSTLSCYGDENKLFAGIRRDG-LLRLVDDFLV 895

Db 748 TVNKKDAKFLNLSRGFEKHNFSTSLKTYVINFENSGIINNTFFNESKKR-M-PFFGF 805

QY 896 TPLHTHAKTFLRVLRGVPEYGCVVNLRKTVVNFPEDEALGGTAFVQMPAHGLFPWGL 955

Db 806 SVNMRSLD 813

QY 956 LLDTRTLE 963

RESULT 6

ID 000939 PRELIMINARY; PRT; 1031 AA.

AC 000939;

DT 01-JUL-1997 (TREMUREL. 04, CREATED)

DT 01-JUL-1997 (TREMUREL. 04, LAST SEQUENCE UPDATE)

DT 01-JAN-1998 (TREMUREL. 05, LAST ANNOTATION UPDATE)

DE TELOMERASE SUBUNIT P123.

OS EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; ALVEOLATA; CILIOPHORA.

OC HYPOTRICHES; EUPLOTIDA; EUPLOTES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 97274210.

RA LINGNER J., HUGHES T.R., SHEVCHENKO A., MANN M., LUNDBLAD V.,

RA CECCH T.R.;

RL SCIENCE 276:561-567 (1997).

DR EMBL; U95964; G2072336; -.

QY SEQUENCE 1031 AA; 122562 MW; 21A885CD CRC32;

Query Match 5.1%; Score 439; DB 5; Length 1031;

Best Local Similarity 23.0%; Pred. No. 1.97e-60;

Matches 142; Conservative 168; Mismatches 261; Indels 47; Gaps 34;

Db 361 FINEFFYNILPKDFLGR-NRKNFQKKVKKYVELNKHE-LIHNKLLLEKINTREISMVQV 418

QY 486 FVRACLRLVPPGLGWSRHRNRRFLNTKFKISLGKHAQLSLOELTW-KMSVRDCAWLR 544

Db 419 ETSAKHFFIYDHE-NIVVLWKLRWIFEDLVLSLIRCFEYVTEQKSYKTYVKNIND 477

QY 545 SPGVCVPAAEHLREILAKFLHLMVSVVVELLSRFFVYVTTTFOKNRLFYRKSVM 604

Db 478 VIMKSI-ADLKKETLAEOKEV-EWKKSLGFAFGKLRIPKTTFRPI--MTF--NK 531

QY 605 KLSIGIRQLHKLVRQLRELSEAEVRQHRARPALLTSRLRFPKPDGLRPVNDYVGA 664

Db 532 KIYNSDRKTKTLTNTKLNLSHMLKTLKNRMKDFGFAVFNDDVMKKEFVCK-WK 590

QY 665 RTFRERKRAERLTSRVKALFS--VLNYERARR-PGLLGASVLGLDDIHRARTFVLVR 721

Db 591 QVGQPKLFFATMDIEKCYDSVNEKLSLTKTKLLSSDFWIMTAOLKKNKIVIDSKN 650

QY 722 QDPPPELYFVKVDVTGAYDTIPODRLTEVI-AS-IIRPQ---NTYCV--RRYAV-Q-KA 772



QY 321 VGRQHGHGPPSTSRPPRPWDTPCPVP 346

RESULT 11

ID 000600 PRELIMINARY; PRT; 234 AA.

AC 000600;

DT 01-JUL-1997 (TREMBLREL. 04, CREATED)

DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)

DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)

GN PAROTID 'O' PROTEIN (FRAGMENT).

GN PR84.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE; 96108375.

RA AZEN E.A., AMBERGER E., FISHER S., PRAKOPHOL A., NIECE R.L.;

RL AM. J. HUM. GENET. 58:143-153(1996).

DR EMBL; S80916; G1911492; -.

FT NON\_TER 1

FT 1

ST SEQUENCE 234 AA; 23656 MW; B5920075 CRC32;

Query Match 1.6%; Score 142; DB 4; Length 234;

Best Local Similarity 28.2%; Pred. No. 4.02e-05;

Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps

Db 92 OGGNOSQ-P-PPHGGKPERPPP-OGG-NOSHRPPPPP-GKPER-PPPGGNSQG-P-PP 144

QY 199 OLGAATQARPPPHASGFRRLGCRANWHSVRAGVPLGLPAPGARRRGGSASRSLPLPK 258

Db 145 HPKKEGPPQEGNKGSRASPPGKPGQPPQOEGNKPQPPPGKPGQPPPGGNNPQ 204

QY 259 RPRGAAPERPTVQGGSWAHGRTGFSDR-GFCVVSFAPAE-EATSLGALSGETH 316

Db 205 APPAGKPGPPPPPQGGRRPPR 226

QY 317 SHPSVGRQHHAGPPSTSRPPR 338

RESULT 12

ID Q46612 PRELIMINARY; PRT; 276 AA.

AC Q46612;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)

DE IS 1222 GENE ORF-A AND ORF-B.

OS ENTEROBACTERIA AGGLOMERANS.

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS

OC ENTEROBACTERIACEAE.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-KLEEGER, 1983;

RX MEDLINE; 95255664.

RA STEIBL H.D., LEWECKE F.M.;

RL GENE 156:37-42(1995).

RN [2]

RP SEQUENCE FROM N.A.

RA STEIBL H.D., SIDDAVATTAM D.;

RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE; 96422758.

RA STEIBL H.D., SIDDAVATTAM D., KLINGMUELLER W.;

RL PLASMID 34:223-228(1995).

RN [4]

RP SEQUENCE FROM N.A.

RA STEIBL H.D.;

RL THESIS (1995), UNIVERSITAET BAYREUTH.

DR EMBL; X78052; G459248; -.

DR EMBL; X81893; E258949; -.

ST SEQUENCE 276 AA; 31718 MW; 0B54A420 CRC32;



Query Match 1.6%; Score 140; DB 2; Length 373;  
Best Local Similarity 32.3%; Pred.No. 7.93e-05;  
Matches 40: Conservative 27; Mismatches 46; Indels 11; Gaps

[illegible]

DB	352	PPRP	352	
		+	+	
QY	335	PPRP	338	
RESULT 15				
ID	041935			P
AC	041935;			
DT	01-JAN-1998			(
DT	01-JAN-1998			(
DT	01-JAN-1998			(
DE	HYPOTHETICAL			
GN	GAMMAHV M6			

OS MURINE HERPESVIRUS 68.  
 RN VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE; GAMMAHERPESVIRINAE  
 [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=WUMS;  
 RC STRAIN=WUMS;  
 RX MEDLINE; 97366649.  
 RA VIRGIN H.W. IV, LATREILLE P., WAMSLEY K., HALLSWORTH K., WECK K.E.

RA	DAL CANTO A. J., SPECK S. H.;
RL	J. VIROL. 71:5894-5904(1997).
RN	[2]
RA	SEQUENCE FROM N. A.
RL	STRAIN=WUMS;
RN	RC
RA	LATRILLE P., WAMSLER P., WATERSTON R. H.;
RL	SUBMITTED (APR-1997) TO EMBL/GENBANK/DDRJ
RN	DATA BANKS.

EMBL: U97553; G2317934; -  
HYPOTHETICAL PROTEIN.  
Q SEQUENCE 585 AA; 60160 MW; B35C72E5 CRC32;

```

Query Match      1.68; Score 135; DB 14; Length 585;
Best Local Similarity 27.6%; Pred. No. 4-23e-04;
Matches 45; Conservative 37; Mismatches 68; Indels 13; Gaps
Db 342 PPP-ELGPGSPTSPASRACARTDPLGGLPWSGPPRPDPPELGGPSPTSPASRAG 400
| | | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | |
QY 195 PPLYQLGAATCAQPPPHASGPRRR-L-GCRAWNHSVREA-GVP-LGLPAPGARRGSA 250

```

QY	251	S R S L P L P K R - P R R G A A P E R T P V Q G S W A H P G R T R G S D R G F C W S E
Db	457	S P R F P R S R P P - - P E L G C G S D P L G P L F S W G - P O P P T F P A D S R 496
QY	310	A L S G T R H S P S V G R O H H A G P P T S R P R P M D T C P P V Y A E T K 352

Search completed: Thu Dec 24 07:59:53 1998  
Job time : 479 secs.

**THIS PAGE BLANK (USPTO)**



```

QY 239 PAPGARRGGASRS-LPLPKRPR-RGAAPERTPVCGSWAHPGRTGSPSD-RGFCVV 295
Db 96 tpppappedgrrpgagnasrgrrsggrrprprpskappkwrk 140
QY 296 SPARPAEATSLGALSCTRHSFSGVGRHAGPPSTSRP-PRDW 339

RESULT 2
ID W18664 standard; Protein; 386 AA.
AC W18664;
DE 24-JUL-1997 (first entry)
KW Framedgment human NF-H gene +1 frameshift mutant product.
KW Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;
KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;
KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;
KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
KW cardiovascular; rheumatoid arthritis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 1..387
FT /note= "X corresponds to a stop codon in the
FT accompanying DNA file, T69796"
FT peptide
FT 129..135
FT /note= "antigenic peptide used fro antibody
FT production"
FT W09712992-A2.
FN 10-APR-1997. PD
PD 02-OCT-1996; IB1106.
PF 02-OCT-1995; GB-020080.
PR 11-JAN-1996; US-009832.
PR (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYUO-) UNIV ROTTERDAM ERASMUS.
PA (UYUT-) UNIV STATE UTRCHTSM.
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
DR WPI; 97-226235/20.
DR N-PSDB; T69796.
PT Use of mutant genes having frame:shift mutation(s) - for developing
PT prods. for the diagnosis, prevention and treatment of associated
PT diseases, e.g. cancer or neuro:degenerative disease
PT Claim 22; Fig 9; 123pp; English.
PS W18663 and W18664 are +2 and +1 frameshift mutations, respectively,
CC of a sequence comprising fragments of the coding sequence of the
CC human neurofilament subunit NF-H gene corresponding to nucleotides
CC 1-1162 of the wild-type NF-H gene. This region contains GAGAG motifs.
CC Frameshift mutants of the tau, ubiquitin, apolipoprotein E,
CC microtubule-associated protein 2 (MAP-2), neurofilament subunit L, M
CC and H and amyloid A4 genes are claimed. All these genes share a common
CC GAGAN motif (N= A, G, C or T), which is the site of common GA
CC dinucleotide deletion(s) that cause neurodegenerative disorders.
CC Antigenic peptides used for the production of antibodies, and small
CC nucleic acid sequences derived from frameshift mutants are used in the
CC diagnosis, prevention and treatment of cancer and neurodegenerative
CC diseases, e.g. Parkinson's disease, Alzheimer's disease, Down's
CC syndrome, frontal lobe dementia (Pick's disease), progressive
CC supranuclear palsy (PSP), amyotrophic lateral sclerosis, Huntington's
CC disease, multiple sclerosis, and other degenerative diseases such as
CC cardiovascular disease and rheumatoid arthritis.
SQ Sequence 386 AA;

Query Match 1.4%; Score 117; DB 22; Length 386;
Best Local Similarity 28.0%; Pred. No. 5.36e+00;
Matches 30; Conservative 26; Mismatches 50; Indels 1; Gaps 1;

Db 228 agglry-graviragpdrargagpgrgatsatgagaparghrapaprrirgpaarg 286
QY 232 AGVFLGLPAGARRGGASRS-LPLPKRPRGAAPERTPVCGSWAHPGRTGSPDRG 291
Db 287 grggpragalragrgrgarpaeegagaagvrlpaappbgrggra 333
QY 292 FCVVSPPARPAEATSLGALSCTRHSFSGVGRHAGPPSTSRPRP 338

```

```

RESULT 3
ID R23846 standard; Protein; 399 AA.
AC R23846;
DE 05-NOV-1992 (first entry)
DE Prepro form of human osteogenic polypeptide-2.
KW Human osteogenic polypeptide-2; prepro sequence; implant;
KW endochondral bone formation; skeletal abnormality; dental abnormality;
KW non-union fracture; cartilage repair; osteoarthritis;
KW periodontal applcn.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cleavage_site 13..18
FT /label= cleavage site for signal peptide removal
FT cleavage_site 257..261
FT /label= cleavage site to form mature protein
FT protein 261..399
FT /label= mature protein 'hOP2-A'
FT protein 264..399
FT /label= truncated protein 'hOP2-P'
FT protein 267..399
FT /label= truncated protein 'hOP2-R'
FT /note= "I"
FT protein 240..399
FT /label= truncated protein 'hOP2-S'
FT region 303..399
FT /label= conserved six cysteine skeleton
FT region 297..399
FT /label= conserved seven cysteine skeleton
FN W09207073-A.
PN 30-APR-1992.
PD 18-OCT-1991; U07635.
PR 18-OCT-1990; US-599543.
PR (CREA-) CREATIVE BIOMOLECULES INC.
PA Kuberassampath T, Oppermann H, Ozkaynak E, Rueger DC;
DR WPI; 92-167153/20.
DR N-PSDB; N24522.
DR New osteogenic polypeptide(s) and their dimeric proteins - useful
PT in association with a matrix for bone reconstruction after
PT orthopaedic and plastic surgery
PT Disclosure; 51-53; 69pp; English.
PS The sequence encodes the prepro form of human osteogenic polypeptide-2
CC (hOP2-PP). The protein is useful as a subunit of osteogenic
CC proteins capable of endochondral bone formation, and allogenic and
CC xenogenic implants in mammals. When implanted with a matrix, the
CC polypeptide locally induces the full developmental cascade of
CC endochondral bone formation and bone marrow differentiation. This
CC can be used for optimal predictable bone formation to correct, eg
CC acquired or congenital craniofacial and other skeletal or dental
CC anomalies, and to treat non-union fractures. The polypeptide also
CC has dental and periodontal applics., and may be used for cartilage
CC repair, eg in osteoarthritis treatment. See also W09105802.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 39.7%; Pred. No. 4.60e+00;
Matches 25; Conservative 13; Mismatches 19; Indels 6; Gaps 6;

Db 12 glalcalggggpgrpp-qc-pqrrlg-ardrdvq-reillavlgpgr-prprappaas 66
QY 194 GPPYQLGANTQA-RPPPHASGPRRLGCCERANWHSYREAGVPLGLPAPARRGGASR 252
Db 67 rlp 69
QY 253 SLP 255

RESULT 4
ID W40193 standard; Protein; 399 AA.
AC W40193;
DE 08-JUN-1998 (first entry)
DE Mouse OP-2 protein.
KW Osteogenic protein; OP-2; bone morphogenic protein; OP/BMP family;
KW chronic renal failure; renal therapeutic agent; disease; diabetes;

```



PR 23-FEB-1989; US-315342.  
 PR 17-OCT-1989; US-422699.  
 PR 17-OCT-1989; US-422613.  
 PR 22-FEB-1990; US-483913.  
 PR 20-AUG-1990; US-569920.  
 PR 07-SEP-1990; US-579865.  
 PR 18-OCT-1990; US-599543.  
 PR 21-NOV-1990; US-616374.  
 PR 04-DEC-1990; US-621849.  
 PR 04-DEC-1990; US-621988.  
 PR 22-FEB-1991; US-660162.  
 PR 20-DEC-1991; US-810560.  
 PR 28-JAN-1992; US-827052.  
 PR 21-FEB-1992; US-841646.  
 PR 01-NOV-1993; US-147023.  
 PR (STYK) STRYKER CORP  
 PI Kuberasampath T, Oppermann H, Ozkaynak E, Rueger DC;  
 DR WPI: 96-010159/01.  
 DR N-PSDB; 702599.  
 DR Antibodies with osteogenic protein binding specificity - used in  
 PT purification of osteogenic proteins, and as antigenic proteins  
 PS Disclosure; Column 139-42; 129pp; English.  
 CC This sequence represents the murine osteogenic protein, mOP-2.  
 CC has homology with proteins in the TGF-beta superfamily. Fragments  
 CC of the mOP-2 protein can be used in the production of dimeric  
 CC peptides which may be used in the generation of antibodies with  
 CC binding specificities for osteogenic proteins. The antibodies are  
 CC capable of binding specifically to an epitope of the osteogenic protein  
 CC and may be used in purification protocols. Osteogenic proteins, such as  
 CC mOP-2, may be used in an implantable osteogenic device which allows  
 CC predictable bone formation to correct acquired and congenital  
 CC craniofacial and other skeletal or dental anomalies. They may be used to  
 CC induce local endochondral bone formation in non-union fractures and in  
 CC other clinical applications including dental and periodontal applications  
 CC where bone formation is required. Other potential applications include  
 CC cartilage repair, e.g. in the treatment of osteoarthritis.  
 CC Sequence 399 AA;  
 SQ

Query Match 1.4%; Score 118; DB 16; Length 399;  
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;  
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
 DB 12 glalcalggghgppp-htc-pqrrlg-arerrdmqreilavlgpr-prprsqaaar 67  
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANVHNSVREAGVPLGLPAPGARRRGSASRS 253  
 Db 68 qp 69  
 QY 254 LP 255

RESULT 7  
 ID R54938 standard; Protein; 399 AA.  
 AC R54938;  
 DT 15-OCT-1994 (first entry)  
 DE Osteogenic protein mOP2-PP.  
 KW Morphogenic protein; mOP-2-PP; OP-2; mOP2; mOP-2;  
 KW Tissue morphogenesis; osteogenic protein.  
 OS Mus sp.  
 PN W09410203-A.  
 PD 11-MAY-1994.  
 PF 02-NOV-1993; U10520.  
 PR 03-NOV-1992; US-971091.  
 PR 04-MAR-1993; US-029335.  
 PR 31-MAR-1993; US-040510.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E;  
 PI Pang RHL, Rueger DC;  
 DR WPI: 94-167392/20.  
 DR P-PSDB; 065394.  
 PT A morphogenically active protein mOP-3 - for inducing tissue  
 PT morphogenesis in mammals  
 PS Disclosure; Page 134-136; 164pp; English.

CC A novel mouse morphogenic protein, OP3, has the sequence given in  
 CC R54934, and is encoded by cDNA of sequence 065390. cDNA and protein  
 CC sequences were also provided for human osteogenic protein OPl  
 CC (Q65391, R54935), mouse OPl (Q65392, R54936), human OP2 (Q65393,  
 CC R54937) and mouse OP2 (Q65394, R54938), as well as the genomic DNA  
 CC sequence of human OP2 (Q65395). Generic sequences given in R54939-  
 CC 40 accommodate homologies between OPl, OP2, OP3 and other morphogen  
 CC protein family members.  
 SQ Sequence 399 AA;  
 Query Match 1.4%; Score 118; DB 10; Length 399;  
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;  
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
 DB 12 glalcalggghgppp-htc-pqrrlg-arerrdmqreilavlgpr-prprsqaaar 67  
 QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANVHNSVREAGVPLGLPAPGARRRGSASRS 253  
 Db 68 qp 69  
 QY 254 LP 255

RESULT 8  
 ID R33908 standard; Protein; 399 AA.  
 AC R33908;  
 DT 21-JUL-1993 (first entry)  
 DE Mouse osteogenic protein 2 (mOP-2).  
 KW Bone; loss; increase; fracture; post-menopausal; senile;  
 KW osteoporosis; hyperparathyroidism; skeletal microstructure defects;  
 KW chronic renal failure; kidney disease; osteomalacia, vitamin D;  
 KW deficiency-induced osteopenia, osteoporosis; Paget's disease;  
 KW bone mass; imbalance; resorption; formation; dialysis; calcium;  
 KW phosphate; metabolism; murine.  
 OS Mus musculus.  
 FH Key  
 FT Location/Qualifiers  
 FT protein 261..399  
 FT region 18..260  
 FT /note= "mature protein"  
 FT /note= "pro region, cleaved to yield mature,  
 FT active protein"  
 FT region 298..299  
 FT /note= "conserved seven cysteine skeleton"  
 FT W09305751-A.  
 PN 01-APR-1993.  
 PF 28-AUG-1992; U07432.  
 PR 30-AUG-1991; US-752764.  
 PR 30-AUG-1991; US-752857.  
 PR 30-AUG-1991; US-752861.  
 PR 31-JUL-1992; US-923780.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E,  
 PI Pang RHL, Rueger DC, Smart JE;  
 DR WPI: 93-117208/14.  
 PT Use of morphogenic or in-vivo morphogenic-stimulating agent - to  
 PT prevent bone loss or increase, used for treating bone fractures,  
 PT post-menopausal or senile osteoporosis, hyperparathyroidism etc.  
 PS Disclosure; Page 122-123; 162pp; English.  
 CC The sequence is that of mouse osteogenic protein 2 (mOP-2) a  
 CC a morphogenically active protein which may be used as part of a  
 CC method for treating a bone fracture or a disease which causes or  
 CC results in bone fractures or other defects in skeletal  
 CC microstructure. Such diseases include chronic renal failure and  
 CC other kidney diseases, osteomalacia, vitamin D deficiency-induced  
 CC osteopenia or osteoporosis, postmenopausal or senile osteoporosis,  
 CC hyperparathyroidism and Paget's disease. The methods can be used for  
 CC protecting individuals at risk for loss of bone mass such as  
 CC postmenopausal females, aged individuals and individuals undergoing  
 CC dialysis. The loss of bone mass may result from an imbalance in bone  
 CC resorption or bone formation, an imbalance of calcium or phosphate  
 CC metabolism, a vitamin D imbalance or be nutritionally or hormonally  
 CC induced.  
 SQ Sequence 399 AA;

KW	tissue.
OS	Mus musculus.
PN	WO9406420-A.
PD	31-MAR-1994.
PF	UO8885.
PR	15-SEP-1993:
PR	15-SEP-1992: US-945286.
PR	04-MAR-1993: US-029335.
PR	31-MAR-1993: US-040510.
PA	(CREA-) CREATIVE BIOMOLECULES INC.
PI	Charrette MF, Cohen CM, Kuberasampath T, Oppermann H;
PI	Ozkanak E, Pang RHL, Rueger DC, Smart JE;
DR	WPI: 94-118121/14.
DR	N-PSDB; Q45119.
PT	Maintaining integrity of gastrointestinal lining using a
PT	morphogen (stimulant) - for treating or preventing ulceration,
PT	also to inhibit endothelial cell proliferation and reduce side
PT	effects of cancer therapy.
PS	Claim 35-36: Page 115-117; 151pp; English.
CC	Morphogens comprising an amino acid sequence sharing at least
CC	70% homology with OP-1, OP-2, CBMP2, BMP3(fx), Vgi(fx), Vgr(fx),
CC	DPP(fx), GDF-1(fx), 60A(fx) and at least 80% homology with
CC	BMP5(fx) and BMP6(fx) are useful for maintaining the integrity of
CC	the gastrointestinal tract luminal lining in a mammal, including
CC	(1) limiting epithelial cell proliferation, (2) inhibiting ulcerative
CC	lesion formation, (3) inhibiting inflammation normally associated
CC	with ulcerative diseases, and/or (4) stimulating the repair of
CC	ulcerative lesions and the regeneration of the luminal tissue.
SQ	Sequence 399 AA;
Query Watch	1.4%; Score 118; DB 10; Length 399;
Best Local Similarity	38.7%; Pred. No. 4.60e+00;
Matches	24; Conservative 12; Mismatches 22; Indels 4; Gaps
Db	12 glialcagghprpp-hcc-bqrrlg-arerrdmqrailavglpgpr-prpracgaaar 67
QY	194 GPPLYQLGATQRPPPHASGPRRRRLGCERANVHSVRKAVGLPGAPCARRGGSAGRS 253
Db	68 qp 69
QY	254 LP 255
RESULT	11
ID	R47253 standard; Protein; 399 AA.
AC	R47253;
DT	15-AUG-1994 (first entry)
DE	MOP2.
KW	Human; hippocampus; osteogenic protein; OP-1; subunit; dimer;
KW	morphogenic activity; cysteine; morphogen; family; pro-region;
KW	complex; soluble; aqueous solvent; therapeutic composition;
KW	symptom-alleviating; co-factor; antibody; diagnosis; assay;
KW	quantitate; mature.
OS	Mus musculus.
PN	WO9403600-A.
PD	17-FEB-1994.
PF	29-JUL-1993: U07189.
PR	31-JUL-1992: US-923780.
PR	04-MAR-1993: US-029335.
PR	31-MAR-1993: US-040510.
PA	(CREA-) CREATIVE BIOMOLECULES INC.
PI	Jones WK, Kuberasampath T, Oppermann H, Ozkanak E;
PI	Rueger DC, Tucker RF;
DR	WPI: 94-065689/08.
DR	N-PSDB; Q56201.
PT	Morphogenic protein soluble complex - for regeneration of tissue
PT	in mammals and diagnosing tissue disorders
PS	Claim 3; Page 75-77; 120pp; English.
CC	This sequence represents the murine derived protein, osteogenic
CC	protein, mop-2. The mature OP-2 protein was used as at
CC	least one subunit in the dimeric protein of the invention. This
CC	dimeric protein comprises a pair of protein subunits which are
CC	associated to give a structure with morphogenic activity. Each
CC	subunit comprises more than 100 amino acids having a pattern of

CC cysteine residues characteristic of the morphogen family. Each  
CC subunit comprises a mature form of a subunit of a member of the  
CC morphogen family, non-covalently complexed with a peptide comprising  
CC a pro-region of a morphogenic family member, to form a complex more  
CC soluble in aqueous solvents than the uncomplexed subunits. The  
CC dimeric protein is useful in a therapeutic composition, pref. also  
CC containing a symptom-alleviating cofactor. The protein and also  
CC corresponding antibody may be used in diagnostic assays, eg. to  
CC quantitate the amount of mature and soluble forms of morphogenic  
CC proteins produced.  
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 9; Length 399;  
Best Local Similarity 38.7%; Pred. No. 4.60e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaar 67  
QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253  
Db 68 qp 69  
QY 254 LP 255

RESULT 12  
ID W00239 standard; Protein; 399 AA.

AC W00239;  
DT 21-NOV-1996 (first entry)  
DE Mouse osteogenic protein OP-2.  
KW Morphogen; Osteogenic protein; dentine; tooth decay; caries;  
KW Morphogenesis; odontoblast; OP-2.  
OS Mus musculus.

FH Key Location/Qualifiers  
FT peptide 1..17  
FT /label= Sig\_peptide  
FT region 18..260  
FT /label= Pro\_region  
FT protein 261..399  
FT /label= Mat\_protein  
FT domain 288..399  
FT /label= 7-Cys\_C-terminal\_domain

PN W09626737-AL.

PD 06-SEP-1996.

PF 14-FEB-1996; U02169.

PR 01-MAR-1995; US-396930.

PA (CREA-) CREATIVE BIOMOLECULES INC.

PI Charette MF, Rutherford RB;

DR WPI; 96-412583/41.

DR N-PSDB; T33444.

PT Use of morphogen(s), e.g. osteogenic proteins, on dentinal surfaces  
PT - for inducing dentine morphogenesis, desensitising teeth or sealing  
PT tooth cavities

PS Disclosure: Page 61-63; 106pp; English.

CC Mouse embryo full-length osteogenic protein OP-2 (W00239)

CC includes a pro-sequence and the morphogenically active mature

CC protein sequence (see also W00224) that includes a 7-Cys C-terminal

CC domain. OP-2 can be expressed from intact or truncated cDNA

CC (T33444) in prokaryotic or eukaryotic host cells. Mature OP-2 and

CC other morphogens (see also W00221-35), partic. human OP-1 (W00221),

CC can be used to induce dentine morphogenesis, to seal dental

CC cavities and to desensitise teeth to pressure and/or temp.

SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 19; Length 399;  
Best Local Similarity 38.7%; Pred. No. 4.60e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaar 67  
QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253  
Db 68 qp 69

QY 254 LP 255

RESULT 13

ID W36856 standard; Protein; 399 AA.

AC W36856; (first entry)

DT 10-MAR-1998

DE Full length sequence of mouse osteogenic protein 2 (hop-2).

KW Mouse osteogenic protein; OP; OP-2; morphogen; morphogenic protein;

KW embryogenesis; organ maintenance; tissue-specific morphogenesis;

KW arthritis; emphysema; osteoporosis; cirrhosis.

OS Mus sp.

FH Key Location/Qualifiers

FT Region 18..260

FT /note= "pro region which is cleaved to yield the

FT mature morphogenically active protein"

FT Protein 261..399

FT /note= "mature protein"

FT Region 297..399

FT /note= "conserved 7 Cys skeleton"

PN US5650276-A.

PD 22-JUL-1997.

PF 20-JUL-1994; 278729.

PR 28-AUG-1992; US-938021.

PR 30-AUG-1991; US-752764.

PR 30-AUG-1991; US-752861.

PA (CREA-) CREATIVE BIOMOLECULES INC.

PI Cohen CM, Kuberasampath T, Oppermann H, Ozkaynak E;

PI Pang RH, Rueger DC, Smart JE;

DR WPI; 97-384865/35.

DR N-PSDB; T97882.

PT Screening for compounds which modulate morphogen expression - by

PT incubating in the presence of epithelial cells which contain a

PT cellular gene for morphogenic protein expression

PT Disclosure; Columns 67-70; 49pp; English.

CC The present sequence represents a mouse osteogenic protein-2 (MOP-2).

CC OP-2 proteins are a group of morphogenically active proteins. Morphogens

CC are inactive when reduced, but are active as oxidised homodimers and when

CC oxidised with other morphogens (e.g. W36853-62). Comparison of the amino

CC acid sequences of these morphogens has identified a consensus 6-7

CC cysteine motif at the C-terminal. Morphogenic proteins such as OP-2 play

CC an important role, not only in embryogenesis, but also in tissue and

CC organ maintenance and repair in mammals. They induce a developmental

CC cascade of tissue-specific morphogenesis in a mammal. A novel method is

CC described for screening a candidate compound for the ability to modulate

CC expression of a cellular gene encoding a naturally occurring morphogenic

CC protein. The candidate compound is incubated with epithelial cells which

CC express the cellular gene, and after a period of time the epithelial

CC cells are assayed for the presence of or the amount of the protein

CC expressed by the cellular gene. A change in the level of the

CC morphogenic protein relative to the level in the epithelial cells in the

CC absence of the candidate compound is indicative of the ability of the

CC compound to modulate expression of the cellular gene. The method can be

CC used to identify compounds which can increase or decrease morphogen

CC production or levels. Such compounds can be used in the treatment of,

CC e.g. arthritis, emphysema, osteoporosis, kidney disease, lung diseases,

CC cardiomyopathy, and cirrhosis of the liver.

SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 26; Length 399;

Best Local Similarity 38.7%; Pred. No. 4.60e+00;

Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalggghgprpp-htc-pqrrlg-arerrdmqrellavlgpgr-prpragpaar 67

QY 194 GPPYLQGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253

Db 68 qp 69

QY 254 LP 255



[illegible]

**THIS PAGE BLANK (USPTO)**

(TM)

Result No.	Score	Query %		Length	DB	ID	Description	Pred. No.
		Match						
1	8624	100.0	1154	13	US-08-951-	Sequence 20, Applicati	0.00e+00	
2	8507	99.8	1189	18	US-08-974-	Sequence 613, Applicat	0.00e+00	
3	8607	99.8	1189	13	US-08-911-	Sequence 34, Applicati	0.00e+00	
4	8607	99.8	1189	13	US-08-912-	Sequence 325, Applicat	0.00e+00	
5	8607	99.8	1200	18	US-08-974-	Sequence 612, Applicat	0.00e+00	
6	8607	99.8	1200	13	US-08-912-	Sequence 324, Applicat	0.00e+00	
7	8607	99.8	1200	13	US-08-911-	Sequence 33, Applicati	0.00e+00	
8	8607	99.8	1285	18	US-08-974-	Sequence 600, Applicat	0.00e+00	
9	8607	99.8	1285	13	US-08-911-	Sequence 32, Applicati	0.00e+00	
10	8607	99.8	1285	13	US-08-912-	Sequence 314, Applicat	0.00e+00	
11	8590	99.6	1407	13	US-08-911-	Sequence 55, Applicati	0.00e+00	
12	8590	99.6	1407	18	US-08-974-	Sequence 628, Applicat	0.00e+00	
13	8590	99.6	1407	13	US-08-912-	Sequence 334, Applicat	0.00e+00	
14	8465	98.2	1132	18	US-08-974-	Sequence 2, Applicatio	0.00e+00	
15	8465	98.2	1132	13	US-08-912-	Sequence 2, Applicatio	0.00e+00	
16	8465	98.2	1132	13	US-08-911-	Sequence 2, Applicatio	0.00e+00	
17	8465	98.2	1132	14	US-09-042-	Sequence 3, Applicatio	0.00e+00	
18	8465	98.2	1132	14	US-09-052-	Sequence 2, Applicatio	0.00e+00	
19	8465	98.2	1132	12	US-08-854-	Sequence 225, Applicat	0.00e+00	
20	8465	98.2	1154	13	US-08-912-	Sequence 323, Applicat	0.00e+00	
21	8465	98.2	1154	18	US-08-974-	Sequence 611, Applicat	0.00e+00	







CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITLE OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US

CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 325:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1189 amino acids  
CC TYPE: amino acid

CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein

CC SEQUENCE 1189 AA; 133179 MW; 7256545 CN;  
Query Match 99.8%; Score 8607; DB 13; Length 1189;  
Best Local Similarity 99.98; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 37 ASTQRCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPPLATFVRRLGPGQGNRLV 96

QY 2 ASQRCVLLRTWEALPATPAMPAPRCRAVRSLLRSHYREVLPPLATFVRRLGPGQGNRLV 61

DB 97 QRGDPAAFRALVAQCLVCVPWDARPPAPPSFRQVSKLKVAVRLQRLCERGAKNYLA 156

QY 62 QRGDPAAFRALVAQCLVCVPWDARPPAPPSFRQVSKLKVAVRLQRLCERGAKNYLA 121

DB 157 GFALLDGARGGPPPEATTTSVRSYLPNTVTDALRGSGANGLLLRVGGDDVLVHLLARCALF 216

|||||

QY 122 GFALLDGARGGPPPEATTTSVRSYLPNTVTDALRGSGANGLLLRVGGDDVLVHLLARCALF 181  
DB 217 VLVAPSCAYQVCPPPLYQLGAATQARPPHASPRRRLGRCERAWNSHREAGVPLGLPAP 276  
QY 182 VLVAPSCAYQVCPPPLYQLGAATQARPPHASPRRRLGRCERAWNSHREAGVPLGLPAP 241  
DB 277 GARRGGASRSRLPLKPRRGAAPERTPVGQGSWAHPGRTRGSDRGFCVVSARPA 336  
QY 242 GARRGGASRSRLPLKPRRGAAPERTPVGQGSWAHPGRTRGSDRGFCVVSARPA 301  
DB 337 EEATSLLEGALSGTRHSHSPVSGROHHAGPSTSRPPRPWDTCCPPVYAEKHFYSSGDKE 396  
QY 302 EEATSLLEGALSGTRHSHSPVSGROHHAGPSTSRPPRPWDTCCPPVYAEKHFYSSGDKE 361  
DB 397 QLRPSLLSLRPSLTGARLVETIFLGRPMWPGTPRRLPRLPQRYWQMRPLFLELLGN 456  
QY 362 QLRPSLLSLRPSLTGARLVETIFLGRPMWPGTPRRLPRLPQRYWQMRPLFLELLGN 421  
DB 457 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSGVAAPPEEDTDPRLVQLLRQHSSPW 516  
QY 422 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSGVAAPPEEDTDPRLVQLLRQHSSPW 481  
DB 517 QYGVFVRACLRRLVPPGLAGSRHNERREFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 576  
QY 482 QYGVFVRACLRRLVPPGLAGSRHNERREFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541  
DB 577 LRSPGVGCVPAEAEHRLREEILAKFLHLMWSVYVVELLSFFVYVTTTFOKNLFFYRKS 636  
QY 542 LRSPGVGCVPAEAEHRLREEILAKFLHLMWSVYVVELLSFFVYVTTTFOKNLFFYRKS 601  
DB 637 VWSKLSIGIRQHLKRVQLRELSAEVROHREARPAALLSRLRFIPKPGDLRPIVNDYV 696  
QY 602 VWSKLSIGIRQHLKRVQLRELSAEVROHREARPAALLSRLRFIPKPGDLRPIVNDYV 661  
DB 697 VGARTERREKRAERLTSRVKALFSVLNYERARPPGLLGASVLGLDDDIHRAWRFTVLVR 756  
QY 662 VGARTERREKRAERLTSRVKALFSVLNYERARPPGLLGASVLGLDDDIHRAWRFTVLVR 721  
DB 757 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVKAF 816  
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKPNQTYCVRRYAVVQKAAHGHVKAF 781  
DB 817 KSHVSTLTLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLFDVFLRPMCHHAVR 876  
QY 782 KSHVSTLTLQPYMROFVAHLQETSPLRDAVTEQSSSLNEASSGLFDVFLRPMCHHAVR 841  
DB 877 IRGKSYVQCGIPQGSILSTLLCSLCYGDMENKLFAGIRBDGLLLRLVDDFLVTPHLTH 936  
QY 842 IRGKSYVQCGIPQGSILSTLLCSLCYGDMENKLFAGIRBDGLLLRLVDDFLVTPHLTH 901  
DB 937 AKTFRLTLVRGVPPEYGCVVNLRKTVVNFVPEDEALGCTAFVQMPAHGLFPWCGLLDTRT 996  
QY 902 AKTFRLTLVRGVPPEYGCVVNLRKTVVNFVPEDEALGCTAFVQMPAHGLFPWCGLLDTRT 961  
DB 997 LEVQSDYSYARTSIRASLTFNRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1056  
QY 962 LEVQSDYSYARTSIRASLTFNRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1021  
DB 1057 NIYKILLQAYREPHACVLOLPFHQOYWKNTPELRYISDTASLCYSILKAKNAGMSLGAK 1116  
QY 1022 NIYKILLQAYREPHACVLOLPFHQOYWKNTPELRYISDTASLCYSILKAKNAGMSLGAK 1081  
DB 1117 GAAGPLPSEAVQWLCHQAFLLKLTNRHRTVYVPLGLSLRTAQTOLSKRLPOTTLTALEAAA 1176  
QY 1082 GAAGPLPSEAVQWLCHQAFLLKLTNRHRTVYVPLGLSLRTAQTOLSKRLPOTTLTALEAAA 1141  
DB 1177 NPALPSDFKTILD 1189  
QY 1142 NPALPSDFKTILD 1154

RESULT 5  
ID US-08-974-549-612 STANDARD; PRT; 1200 AA.

XX  
AC  
XX  
DT  
XX  
XX  
DE  
XX  
XX  
Sequence 612, Application US/08974549  
XX  
Sequence 612, Application US/08974549  
CC  
GENERAL INFORMATION:  
CC  
APPLICANT: Cech, Thomas R.  
CC  
APPLICANT: Lingner, Joachim  
CC  
APPLICANT: Nakamura, Toru  
CC  
APPLICANT: Chapman, Karen B.  
CC  
APPLICANT: Morin, Gregg B.  
CC  
APPLICANT: Harley, Calvin B.  
CC  
APPLICANT: Andrews, William H.  
CC  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
CC  
NUMBER OF SEQUENCES: 726  
CC  
CORRESPONDENCE ADDRESS:  
CC  
ADDRESSEE: Townsend and Townsend and Crew LLP  
CC  
STREET: Two Embarcadero Center, Eighth Floor  
CC  
CITY: San Francisco  
CC  
STATE: California  
CC  
COUNTRY: USA  
CC  
ZIP: 94111-3834  
CC  
COMPUTER READABLE FORM:  
CC  
MEDIUM TYPE: Floppy disk  
CC  
COMPUTER: IBM PC compatible  
CC  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CC  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC  
CURRENT APPLICATION DATA:  
CC  
APPLICATION NUMBER: US/08/974,549  
CC  
FILING DATE: 19-NOV-1997  
CC  
CLASSIFICATION: 536  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/724,643  
CC  
FILING DATE: 01-OCT-1996  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/844,419  
CC  
FILING DATE: 18-APR-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/846,017  
CC  
FILING DATE: 25-APR-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/851,843  
CC  
FILING DATE: 06-MAY-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/854,050  
CC  
FILING DATE: 09-MAY-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/911,312  
CC  
FILING DATE: 14-AUG-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/912,951  
CC  
FILING DATE: 14-AUG-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 08/915,503  
CC  
FILING DATE: 14-AUG-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: WO PCT/US97/17618  
CC  
FILING DATE: 01-OCT-1997  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: WO PCT/US97/17885  
CC  
FILING DATE: 01-OCT-1997  
CC  
ATTORNEY/AGENT INFORMATION:  
CC  
NAME: Apple, Randolph Ted  
CC  
REGISTRATION NUMBER: 36,429  
CC  
REFERENCE/DOCKET NUMBER: 015389-0026100S  
CC  
TELECOMMUNICATION INFORMATION:  
CC  
TELEPHONE: (415) 576-0200  
CC  
TELEFAX: (415) 576-0300  
CC  
INFORMATION FOR SEQ ID NO: 612:  
CC  
SEQUENCE CHARACTERISTICS:

CC  
LENGTH: 1200 amino acids  
CC  
TYPE: amino acid  
CC  
STRANDEDNESS:  
CC  
TOPOLOGY: linear  
CC  
MOLECULE TYPE: protein  
CC  
FEATURE:  
CC  
NAME/KEY: Protein  
CC  
LOCATION: 1..1200  
CC  
OTHER INFORMATION: /note= "fusion protein composed of His6  
CC  
OTHER INFORMATION: and Anti-Xpress tags, enterokinase  
CC  
OTHER INFORMATION: cleavage site and full length HTRT  
CC  
OTHER INFORMATION: protein"  
CC  
SEQUENCE 1200 AA; 134322 MW; 7387257 CN;  
Query Match 99.8%; Score 8607; DB 18; Length 1200;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 48 ASTQRCVLLTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRWLV 107  
QY 2 ASQRCVLLTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQGRWLV 61  
Db 108 QRGDPAAFRALVAQCLVCPWDARPPAPSPFQVSCLELVARVLRQRCERGAKNVLA 167  
QY 62 QRGDPAAFRALVAQCLVCPWDARPPAPSPFQVSCLELVARVLRQRCERGAKNVLA 121  
Db 168 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGALLLRVGGDDVLVHLLARCALF 227  
QY 122 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGALLLRVGGDDVLVHLLARCALF 181  
Db 228 VLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRLRCERAWNSHVREAGVPLGLPAP 287  
QY 182 VLVAPSCAYQVCGPPLYQLGAATQARPPPHASGPRRLRCERAWNSHVREAGVPLGLPAP 241  
Db 288 GARRRGGSASRLPLKPRRGAAPERTPVQGSWAHPGRTGRGSCFVVSARPA 347  
QY 242 GARRRGGSASRLPLKPRRGAAPERTPVQGSWAHPGRTGRGSCFVVSARPA 301  
Db 348 EEATSLLEGALSGTRHSHSPSVGROHAGPSTSRPPPMWDTCPVYAEAKHFLYSSGDKE 407  
QY 302 EEATSLLEGALSGTRHSHSPSVGROHAGPSTSRPPPMWDTCPVYAEAKHFLYSSGDKE 361  
Db 408 QLRSFLLSRLPSLTGARLVETIFLGSRPWMPGTPRRLPLRPQRYWQMRPLFLELLGN 467  
QY 362 QLRSFLLSRLPSLTGARLVETIFLGSRPWMPGTPRRLPLRPQRYWQMRPLFLELLGN 421  
Db 468 HAOCYPGVLLKTHCPRLAAVTPAAGVCAREKPGQGSVAAPPEEDTDPRLLVQLLRQHSPP 527  
QY 422 HAOCYPGVLLKTHCPRLAAVTPAAGVCAREKPGQGSVAAPPEEDTDPRLLVQLLRQHSPP 481  
Db 528 QYGFVRACLRLRVLPPGLWGSRHNRERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 587  
QY 482 QYGFVRACLRLRVLPPGLWGSRHNRERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541  
Db 588 LRSPGVGCVPAEAEHLRBEILAKFLHMLMSVYVVELLSFFVYTTTQKNRLLFFYRKS 647  
QY 542 LRSPGVGCVPAEAEHLRBEILAKFLHMLMSVYVVELLSFFVYTTTQKNRLLFFYRKS 601  
Db 648 VMSKLSQIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNDIV 707  
QY 602 VMSKLSQIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPGCLRPVNDIV 661  
Db 708 VGARTFRREKRAERLT SRVKALF SVLNERARRPGLLGASVLGLDDIHRARTFVLVRVA 767  
QY 662 VGARTFRREKRAERLT SRVKALF SVLNERARRPGLLGASVLGLDDIHRARTFVLVRVA 721  
Db 768 QDPPPELYFKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVRKAF 827  
QY 722 QDPPPELYFKVDVTGAYDTIPQDRLTEVIASIKPONTYCVRRYAVVQKAAHGHVRKAF 781  
Db 828 KSHVSTLTDLQPYMROFVAHLQETSPLRDANVIEQSSSLNEASSGLDFVLRPMCHHAVR 887  
QY 782 KSHVSTLTDLQPYMROFVAHLQETSPLRDANVIEQSSSLNEASSGLDFVLRPMCHHAVR 841



Db 888 IRKSVVQOGIPQGSILSTLLCSLCYGDENKLFAGIRDRGGLLLRLVDDFLVTPHLTH 947  
QY 842 IRKSVVQOGIPQGSILSTLLCSLCYGDENKLFAGIRDRGGLLLRLVDDFLVTPHLTH 901  
Db 948 AKTFRLTVRGVPEYGVNLRKTVNFPVEDEALGTAFOVQMPAHGLFPWCGLLDTRT 1007  
QY 902 AKTFRLTVRGVPEYGVNLRKTVNFPVEDEALGTAFOVQMPAHGLFPWCGLLDTRT 961  
Db 1008 LEVQSDYSYARTSIRASLTFNRGFKAGRNMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1067  
QY 962 LEVQSDYSYARTSIRASLTFNRGFKAGRNMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1021  
Db 1068 NIKYKILLQAYRHACVQLPFFHQVQWKNPTFFLRVISTDASLCYSILKAKNAGMSLGAK 1127  
QY 1022 NIKYKILLQAYRHACVQLPFFHQVQWKNPTFFLRVISTDASLCYSILKAKNAGMSLGAK 1081  
Db 1128 GAAGPLPSEAVONLCHOAFLLKLTNRHVTVPVLLGSLRTAQTOLSKLPCTTTLTALEAAA 1187  
QY 1082 GAAGPLPSEAVONLCHOAFLLKLTNRHVTVPVLLGSLRTAQTOLSKLPCTTTLTALEAAA 1141  
Db 1188 NPALPSDFKTILD 1200  
QY 1142 NPALPSDFKTILD 1154

RESULT 6  
ID US-08-912-951-324 STANDARD; PRT; 1200 AA.

XX xxxxxx

Sequence 324, Application US/08912951

Sequence 324, Application US/08912951

GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS  
NUMBER OF SEQUENCES: 335  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 324:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1200 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1200 AA; 134322 MW; 7387257 CN;  
Query Match 99.8%; Score 8607; DB 13; Length 1200;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 48 ASTQRCVLLRTWEALAPATPAMPRAPCRAVRSLRSLSHRYEVLPATFVRRRLGPOGWRLV 107  
QY 2 ASGQRCVLLRTWEALAPATPAMPRAPCRAVRSLRSLSHRYEVLPATFVRRRLGPOGWRLV 61  
Db 108 QRGDPAAFALVAQCLVCPWDARPPAPSPROVSLKELVARVLRQRCERGAKNVLA 167  
QY 62 QRGDPAAFALVAQCLVCPWDARPPAPSPROVSLKELVARVLRQRCERGAKNVLA 121  
Db 168 GFALLDGGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVDDVLLHLLARCALF 227  
QY 122 GFALLDGGARGGPEAFTTSVRSYLPNTVDALRGSGAWGLLRVDDVLLHLLARCALF 181  
Db 228 VLVAPSCAYOVCGPPLYQLGAATQARPPHAGSPRRRLCERAWNHSVREAGVPLGLPAP 287  
QY 182 VLVAPSCAYOVCGPPLYQLGAATQARPPHAGSPRRRLCERAWNHSVREAGVPLGLPAP 241  
Db 288 GARRGGASRSLLPKRPRRGAAPERTPVQGSWAHPGTRGSDRGFCVVSAPRA 347  
QY 242 GARRGGASRSLLPKRPRRGAAPERTPVQGSWAHPGTRGSDRGFCVVSAPRA 301  
Db 348 EEATSLLEGALSGTRHSHPSVGRHSHAGPPSTSRPPRPWDTPCPVYAEKHFVLYSSGDKE 407  
QY 302 EEATSLLEGALSGTRHSHPSVGRHSHAGPPSTSRPPRPWDTPCPVYAEKHFVLYSSGDKE 361  
Db 408 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPLRPLQRYWQMRPLFLELLGN 467  
QY 362 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPLRPLQRYWQMRPLFLELLGN 421  
Db 468 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAPEEEDTDPRLVQLLRQHSPPW 527  
QY 422 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVAPEEEDTDPRLVQLLRQHSPPW 481  
Db 528 QVYGFVRACLRRLVPPGLGSRHNRERLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 587  
QY 482 QVYGFVRACLRRLVPPGLGSRHNRERLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541  
Db 588 LRRSPGVGCVPAAEHRLREEILAKFLHLMMSVYVVELLSFFVYTTTFTQKNRLLFFYRKS 647  
QY 542 LRRSPGVGCVPAAEHRLREEILAKFLHLMMSVYVVELLSFFVYTTTFTQKNRLLFFYRKS 601  
Db 648 VWSKLQSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSLRFLIPKPDGLRPIVNDYV 707  
QY 602 VWSKLQSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSLRFLIPKPDGLRPIVNDYV 661

Db 708 VGATFFREKRAELTSRVKALFSLVLYNERARRPGLGASVGLGDDTHRAWRFTVLVRA 767  
QY 662 VGATFFREKRAELTSRVKALFSLVLYNERARRPGLGASVGLGDDTHRAWRFTVLVRA 721  
Db 768 ODPPPELYFKVDVTGAYDIPQDRLTEVIASIIKPNQTYCVRRYAVVQAAHGHVRKAF 827  
QY 722 ODPPPELYFKVDVTGAYDIPQDRLTEVIASIIKPNQTYCVRRYAVVQAAHGHVRKAF 781  
Db 828 KSHVSTLTDLPYMRQFVAHLQETSPRLDAVVIQSSSLNEASSGLDFVLFWMCHHAVR 887  
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPRLDAVVIQSSSLNEASSGLDFVLFWMCHHAVR 841  
Db 888 IRGKSYVQCGIPOGSIILSLCSLCYGDMEKLFAGIRGDLRLVDDFLVTPHLTH 947  
QY 842 IRGKSYVQCGIPOGSIILSLCSLCYGDMEKLFAGIRGDLRLVDDFLVTPHLTH 901  
Db 948 AKTFLRTLVRGVPYGCVNLRKTVNPNVEDEALGTAFAVQMPAHGLFPWCGLLDTRT 1007  
QY 902 AKTFLRTLVRGVPYGCVNLRKTVNPNVEDEALGTAFAVQMPAHGLFPWCGLLDTRT 961  
Db 1008 LEVQSDYSYARTSIRASLTFNRGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1067  
QY 962 LEVQSDYSYARTSIRASLTFNRGFKAGRNRRKLFGLVRLKCHSLFLDLQVNSLQTVCT 1021  
Db 1068 NIYKILLQAYRFHACVQLPFFHQOVWKNPTFFLRVVISDTSALCYSLKAKNAGMSLGAK 1127  
QY 1022 NIYKILLQAYRFHACVQLPFFHQOVWKNPTFFLRVVISDTSALCYSLKAKNAGMSLGAK 1081  
Db 1128 GAAGPLPSEAVOMLCHOAFLLKTLRHRVTYVPLIGSLRTAQTQLSRKLPGTTTILEAAA 1187  
QY 1082 GAAGPLPSEAVOMLCHOAFLLKTLRHRVTYVPLIGSLRTAQTQLSRKLPGTTTILEAAA 1141  
Db 1188 NPALPSDFKTILD 1200  
QY 1142 NPALPSDFKTILD 1154

RESULT 7  
ID US-08-911-312-33 STANDARD; PRT: 1200 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX

Sequence 33, Application US/08911312

Sequence 33, Application US/08911312

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William

TITLE OF INVENTION: Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/911,312

FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-002500US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 33:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1200 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1200 AA; 134322 MW; 7387257 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1200;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 48 ASTQRCVLLTWEALAPATPAMPAPRCRAVRSLRSHRYEVLPLATFYRRRLGPGQWRVLV 107  
QY 2 ASQRCVLLTWEALAPATPAMPAPRCRAVRSLRSHRYEVLPLATFYRRRLGPGQWRVLV 61  
Db 108 QRGDPAAFRALVAQCLVCPWDARPPAPSPRQVSLKELVARVLQRLCERGAKNVLA 167  
QY 62 QRGDPAAFRALVAQCLVCPWDARPPAPSPRQVSLKELVARVLQRLCERGAKNVLA 121  
Db 168 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALF 227  
QY 122 GFALLDGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGGDDVLVHLLARCALF 181  
Db 228 VLVAPSCAYQVCGPPLYQLGAATQARPPHASPGRRLRCERAWNHSVREAGVPLGLPAP 287  
QY 182 VLVAPSCAYQVCGPPLYQLGAATQARPPHASPGRRLRCERAWNHSVREAGVPLGLPAP 241  
Db 288 GARRRGGSASRSILPLKPRPRGAAPPEPTPVQGSWAHPGTRGSDRGFCVSPARPA 347  
QY 242 GARRRGGSASRSILPLKPRPRGAAPPEPTPVQGSWAHPGTRGSDRGFCVSPARPA 301  
Db 348 EEAATSLGALSGTRSHSPVGRQHAGPSTSRPPPPMDTPCPVYAEKHFLLYSSGDKE 407  
QY 302 EEAATSLGALSGTRSHSPVGRQHAGPSTSRPPPPMDTPCPVYAEKHFLLYSSGDKE 361  
Db 408 QLRPSFLLSSRLPSLTGARRLVETIFLGSRPWNMPGTTPRRLPRLPQRYQWMLFLELLGN 467  
QY 362 QLRPSFLLSSRLPSLTGARRLVETIFLGSRPWNMPGTTPRRLPRLPQRYQWMLFLELLGN 421  
Db 468 HAQCPYGVLLKTHCPRLAAVTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSPPW 527  
QY 422 HAQCPYGVLLKTHCPRLAAVTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSPPW 481

Db 528 QVYGFVACLRRLVPPGLWGRHNERFLRNTKFKISLGHAKLSLOELTWKMSVRDCAW 587  
QY 482 QVYGFVACLRRLVPPGLWGRHNERFLRNTKFKISLGHAKLSLOELTWKMSVRDCAW 541  
Db 588 LRSPGVGCVPAAEHRLREILAKFLHMLSVVVELLSFEFFVYTTTFKKNRLFYFRKS 647  
QY 542 LRSPGVGCVPAAEHRLREILAKFLHMLSVVVELLSFEFFVYTTTFKKNRLFYFRKS 601  
Db 648 VWSKLSIGIRQLKRVQLRELSEAEVQHQREARPAALLTSRLRFIPKPDGLRPIVNMYY 707  
QY 602 VWSKLSIGIRQLKRVQLRELSEAEVQHQREARPAALLTSRLRFIPKPDGLRPIVNMYY 661  
Db 708 VGARTFRERKRAELTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 767  
QY 662 VGARTFRERKRAELTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 721  
Db 768 QDPPPELYFVKVDVTGAYDIPDORLTEVIASIKPONTYCVRRYAVVQKAAGHVKAF 827  
QY 722 QDPPPELYFVKVDVTGAYDIPDORLTEVIASIKPONTYCVRRYAVVQKAAGHVKAF 781  
Db 828 KSHVSTLTDLPYMRQFVAHLQETSPDRDAVVEQSSSLNEASSGLDFVFLRFMCHHAVR 887  
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPDRDAVVEQSSSLNEASSGLDFVFLRFMCHHAVR 841  
Db 888 IRKSVYQCOGIPQGSILSTLLSLCYGDMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 947  
QY 842 IRKSVYQCOGIPQGSILSTLLSLCYGDMENKLFAGIRRDGLLLRLVDDFLVTPHLTH 901  
Db 948 AKTFRLTVRGVPEYGCNNLRKTVNFPVEDEALGTAFTVQMPAHGLFPWCGLLDTRT 1007  
QY 902 AKTFRLTVRGVPEYGCNNLRKTVNFPVEDEALGTAFTVQMPAHGLFPWCGLLDTRT 961  
Db 1008 LEVQSDYSSVARTSIRASLTFRNGFRAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTVCT 1067  
QY 962 LEVQSDYSSVARTSIRASLTFRNGFRAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTVCT 1021  
Db 1068 NIKYKILLQAYRHACVQLQPFHQVWKNPTFFELRVISDTSASLCYSILKAKNAGMSLGAK 1127  
QY 1022 NIKYKILLQAYRHACVQLQPFHQVWKNPTFFELRVISDTSASLCYSILKAKNAGMSLGAK 1081  
Db 1128 GAAGPLPSEAVOMLCHQAFLLKTLRHRVTYVPLGLSLRTAQTQLSRKLPGTTTILEAAA 1187  
QY 1082 GAAGPLPSEAVOMLCHQAFLLKTLRHRVTYVPLGLSLRTAQTQLSRKLPGTTTILEAAA 1141  
Db 1188 NPALPSDFKTILD 1200  
QY 1142 NPALPSDFKTILD 1154

RESULT 8  
ID US-08-974-549-600 STANDARD; PRT; 1285 AA.

XX XXXXX

Sequence 600, Application US/08974549

Sequence 600, Application US/08974549

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 726

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/974,549  
CC FILING DATE: 19-NOV-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph Ted  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-0026100S  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 600:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1285 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: Protein  
CC LOCATION: 1..1285  
CC OTHER INFORMATION: /note= "fusion protein composed of  
CC OTHER INFORMATION: enterokinase cleavable, His tagged  
CC OTHER INFORMATION: thioredoxin moiety and full length hTERT"  
CC SEQUENCE 1285 AA; 143529 MW; 849280 CN;

Query Match 99.8%; Score 8607; DB 18; Length 1285;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 133 ASTQRCVLLRTWEALAPATPAMPAPRCRAVRSLLSHREVLPATFVRRLGPGQWRV 192

QY 2 ASQRCVLLRTWEALAPATPAMPAPRCRAVRSLLSHREVLPATFVRRLGPGQWRV 61

Db 193 ORGDPAAFRALVAQCLVCPWDARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVIAF 252  
QY 62 ORGDPAAFRALVAQCLVCPWDARPPPAAPSFQVSCCLKELVARVLQRLCERGAKNVIAF 121  
Db 253 GFALLDARGGPPPAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLVHLLARCALF 312  
QY 122 GFALLDARGGPPPAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLVHLLARCALF 181  
Db 313 VLVAPSCAYQVCGPPLYLQGAATQARPPPHASGPRRLGCRERAWNHSVREAGVPLGLPAP 372  
QY 182 VLVAPSCAYQVCGPPLYLQGAATQARPPPHASGPRRLGCRERAWNHSVREAGVPLGLPAP 241  
Db 373 GARRGGSASRLPLPRPRRGAAPERTPVGQSWAHFPGTRGPDGRCFVVSVPARPA 432  
QY 242 GARRGGSASRLPLPRPRRGAAPERTPVGQSWAHFPGTRGPDGRCFVVSVPARPA 301  
Db 433 EATSLGALSGLTRHSHPSVGRQHAGPPSTSRPRPDWTPCPVYAETHFLYSSGDKE 492  
QY 302 EATSLGALSGLTRHSHPSVGRQHAGPPSTSRPRPDWTPCPVYAETHFLYSSGDKE 361  
Db 493 QLRPSFLLSLRLPSLTGARRLVETIFLGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGN 552  
QY 362 QLRPSFLLSLRLPSLTGARRLVETIFLGSRPMPGTPRRPLRPLQRYWQMRPLFLELLGN 421  
Db 553 HAQCPYGVLLKTHCPRLAAVTTPAAGVCAREKPOGQSVAAPEEDTDPRLVOLLRQHSSPW 612  
QY 422 HAQCPYGVLLKTHCPRLAAVTTPAAGVCAREKPOGQSVAAPEEDTDPRLVOLLRQHSSPW 481  
Db 613 QVYGFVACLRRLVPPGLWGRHNRERFLNTRKFIISLGHAKLSLOELTWKMSVRCAW 672  
QY 482 QVYGFVACLRRLVPPGLWGRHNRERFLNTRKFIISLGHAKLSLOELTWKMSVRCAW 541  
Db 673 LRRSPGVCGVPAAEHRLREELIAKELHMLSVYVVELLSFFVYTTTFQKNRLLFFYRKS 732  
QY 542 LRRSPGVCGVPAAEHRLREELIAKELHMLSVYVVELLSFFVYTTTFQKNRLLFFYRKS 601  
Db 733 VWSKLQSIGIRQHLLKRVQLRELSAEVQROHREARPAITSLRIPRIPDGLRPIVNDYV 792  
QY 602 VWSKLQSIGIRQHLLKRVQLRELSAEVQROHREARPAITSLRIPRIPDGLRPIVNDYV 661  
Db 793 VGARTFRERKRAERLTSRVKALFVLNRYERARRPGLLGASVGLGDDTHRAWRFTVLVRA 852  
QY 562 VGARTFRERKRAERLTSRVKALFVLNRYERARRPGLLGASVGLGDDTHRAWRFTVLVRA 721  
Db 853 QDPPPELYFVKVDVTGAYDTIPQDRLEFVIAIIKPNQTCVRRYAVVQKAAHGHVRKAF 912  
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLEFVIAIIKPNQTCVRRYAVVQKAAHGHVRKAF 781  
Db 913 KSHVSTLTDLPYMRQFVAHLQETSPLRDQVVIQSSSLNEASSGLDFVFLRFMCHHAVR 972  
QY 782 KSHVSTLTDLPYMRQFVAHLQETSPLRDQVVIQSSSLNEASSGLDFVFLRFMCHHAVR 841  
Db 973 IRGKSYVQCOGIPQGSILSTLCSLCYGDMEKLFAGIRRDGLLLRLVDDFLVTPHLTH 1032  
QY 842 IRGKSYVQCOGIPQGSILSTLCSLCYGDMEKLFAGIRRDGLLLRLVDDFLVTPHLTH 901  
Db 1033 AKTELRLTVRGVPEYGCVNLRKTVVNFVPEDEALGCTAFVQMPAHGLFPWCGLLLDTRT 1092  
QY 902 AKTELRLTVRGVPEYGCVNLRKTVVNFVPEDEALGCTAFVQMPAHGLFPWCGLLLDTRT 961  
Db 1093 LEVQSDYSSVARTSIRASLTFNRGFKAGRNMRKLFVRLKCHSLFDLDQVNSLQVCT 1152  
QY 962 LEVQSDYSSVARTSIRASLTFNRGFKAGRNMRKLFVRLKCHSLFDLDQVNSLQVCT 1021  
Db 1153 NIYKILLQAVRFHACVLOLPFHQOVWKNPTFFLRVTSDFASLCSYTLKAKNAGMSLGAK 1212  
QY 1022 NIYKILLQAVRFHACVLOLPFHQOVWKNPTFFLRVTSDFASLCSYTLKAKNAGMSLGAK 1081  
Db 1213 GAAGPLPSEAVQWLVCHQAFLLKLTFRHRTVTVPLGLSLRFTAQTLQSRKLPGLTTLTAAEA 1272  
QY 1082 GAAGPLPSEAVQWLVCHQAFLLKLTFRHRTVTVPLGLSLRFTAQTLQSRKLPGLTTLTAAEA 1141  
Db 1273 NPALPSDFKILTD 1285

QY 1142 NPALPSDFKILTD 1154

RESULT 9  
ID US-08-911-312-32 STANDARD; PRT; 1285 AA.  
XX XXXXXX  
AC  
XX  
DT  
XX  
DE  
XX  
Sequence 32, Application US/08911312  
Sequence 32, Application US/08911312  
GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911.312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-00250005  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 32:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1285 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:

CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1285 AA; 143529 MW; 8449280 CN;  
  
Query Match 99.8%; Score 8607; DB 13; Length 1285;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Db 133 ASTQRCVLLTWTALPATAMPAPRCRAVRSLRLSHYREVLPATFVRRLGPGQWRV 192  
QY 2 ASQRCVLLTWTALPATAMPAPRCRAVRSLRLSHYREVLPATFVRRLGPGQWRV 61  
  
Db 193 QRGDPAAFRALVAQCLVCVPWDARPPAAPSFQVSCCLKELVARVLRQLCERGAKNVLA 252  
QY 62 QRGDPAAFRALVAQCLVCVPWDARPPAAPSFQVSCCLKELVARVLRQLCERGAKNVLA 121  
  
Db 253 GFALLDGARGPPEAFETTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLHLLARCALF 312  
QY 122 GFALLDGARGPPEAFETTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLHLLARCALF 181  
  
Db 313 VLVAPSCAYQVCGPPLYLQGAATQARPPPHASGPRRLRCGERAWNHSVRAGVPLGLPAP 372  
QY 182 VLVAPSCAYQVCGPPLYLQGAATQARPPPHASGPRRLRCGERAWNHSVRAGVPLGLPAP 241  
  
Db 373 GARRGGASRSPLPKRPRRGAPEPTPVGOGSWAHGPRTRGSDRGFCVVSAPRA 432  
QY 242 GARRGGASRSPLPKRPRRGAPEPTPVGOGSWAHGPRTRGSDRGFCVVSAPRA 301  
  
Db 433 EEATSLGALSGTRHSHPSVGRQHAGPSTSRPPRPWDTPCPVPVTAETHFLYSSGDKE 492  
QY 302 EEATSLGALSGTRHSHPSVGRQHAGPSTSRPPRPWDTPCPVPVTAETHFLYSSGDKE 361  
  
Db 493 QLRSFLLSLRLSPTGARLVETIFGSRPMPGTPRRLPRLPQRYQWMPRLFLELLGN 552  
QY 362 QLRSFLLSLRLSPTGARLVETIFGSRPMPGTPRRLPRLPQRYQWMPRLFLELLGN 421  
  
Db 553 HAQCPYGLKTKHCPRAAYTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLRQHSPP 612  
QY 422 HAQCPYGLKTKHCPRAAYTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLRQHSPP 481  
  
Db 613 QVTGFVACLRRLVPPGLGWSRHNREFLNTRKFFISLGHAKLSLQELTWKMSVRDCAW 672  
QY 482 QVTGFVACLRRLVPPGLGWSRHNREFLNTRKFFISLGHAKLSLQELTWKMSVRDCAW 541  
  
Db 673 LRSPGVCVPAEHRUREILAKFLHLMVSVYVELLSFFYYTTFQKNRLLFFYRKS 732  
QY 542 LRSPGVCVPAEHRUREILAKFLHLMVSVYVELLSFFYYTTFQKNRLLFFYRKS 601  
  
Db 733 VWSKLSIGIRHLLKRVQLRELSEAEVQROHREAPALLTSRLRFPKPDGLRPIVNDYV 792  
QY 602 VWSKLSIGIRHLLKRVQLRELSEAEVQROHREAPALLTSRLRFPKPDGLRPIVNDYV 661  
  
Db 793 VGARTFREKRAERLTSRKALFSLVANYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 852  
QY 562 VGARTFREKRAERLTSRKALFSLVANYERARRPGLLGASVLGLDDIHRARWTFVLVRRA 721  
  
Db 853 QDPPPELYFVKVDVTGAYDTIPQDLREVIASIIKQNTYCVRRYAYVQKAAHGHVKA 912  
QY 722 QDPPPELYFVKVDVTGAYDTIPQDLREVIASIIKQNTYCVRRYAYVQKAAHGHVKA 781  
  
Db 913 KSHVSTLTDLPQYMRQFVAHLQETSPRDVAVIQSSSLNEASGLDFVLRFWCHAVR 972  
QY 782 KSHVSTLTDLPQYMRQFVAHLQETSPRDVAVIQSSSLNEASGLDFVLRFWCHAVR 841  
  
Db 973 IRGKSYVQCOGIPGSGTSLTSLCISLQYGDMDENKLFAGIRRDGLLLRLVDDFLVTPHLTH 1032  
QY 842 IRGKSYVQCOGIPGSGTSLTSLCISLQYGDMDENKLFAGIRRDGLLLRLVDDFLVTPHLTH 901  
  
Db 1033 AKTFRLTLVRGVEYGCWNLRTKTVVNFVDEALGCTAFVQMPAHGLFPWCGLLDTRT 1092  
QY 902 AKTFRLTLVRGVEYGCWNLRTKTVVNFVDEALGCTAFVQMPAHGLFPWCGLLDTRT 961  
  
Db 1093 LEVQSDYSSYARTSIRASLTFNRGFKAGRNMRKLFQVLRKCHKSLFLDLQVNSLQTVCT 1152

QY 962 LEVQSDYSSYARTSIRASLTFNRGFKAGRNMRKLFQVLRKCHKSLFLDLQVNSLQTVCT 1021  
Db 1153 NIYKILLQAYRFHACVQLQPFHQQVWKKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAK 1212  
QY 1022 NIYKILLQAYRFHACVQLQPFHQQVWKKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAK 1081  
Db 1213 GAAGPLPSEAVOWLCHQAFLLKLTNRHTVYVPLLSRLTAQTSLSRKLPGTTLTALEAAA 1272  
QY 1082 GAAGPLPSEAVOWLCHQAFLLKLTNRHTVYVPLLSRLTAQTSLSRKLPGTTLTALEAAA 1141  
Db 1273 NPALPSDFKTILD 1285  
QY 1142 NPALPSDFKTILD 1154  
  
RESULT 10  
ID US-08-912-951-314 STANDARD; PRT; 1285 AA.  
AC xxxxxx  
DT  
DT  
XX  
XX  
DE Sequence 314, Application US/08912951  
XX  
CC Sequence 314, Application US/08912951  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITL OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-00260005  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 314:  
CC .SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1285 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1285 AA; 143529 MW; 8449280 CN;

Query Match 99.8%; Score 8607; DB 13; Length 1285;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 133 ASTORCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRLPGQGRWLV 192  
|||  
Qy 2 ASGQRCVLLRTWEALAPATPAMPAPRCRAVRSLLRSHYREVLPATFVRLPGQGRWLV 61  
|||  
Db 193 ORGDPAAFRALVACLVCPMDARPPAPAFSROVSCCLKELVARYLQRCERGAKNVLA 252  
|||  
Qy 62 ORGDPAAFRALVACLVCPMDARPPAPAFSROVSCCLKELVARYLQRCERGAKNVLA 121  
|||  
Db 253 GFALLDGARGPPPEAFITSVRSYLPNTVTDALRGSGAWGLLRVGDVLLVHLLARCALF 312  
|||  
Qy 122 GFALLDGARGPPPEAFITSVRSYLPNTVTDALRGSGAWGLLRVGDVLLVHLLARCALF 181  
|||  
Db 313 VLVAPSCAYQVCGPPLVQLGAATQARPAPHASGPRRLGGERAWNHVSREAGVPLGLPAP 372  
|||  
Qy 182 VLVAPSCAYQVCGPPLVQLGAATQARPAPHASGPRRLGGERAWNHVSREAGVPLGLPAP 241  
|||  
Db 373 GARRGGASRLPLKRPREGAPEPTPVGGSWAHGRTGPGDRGFCVVSARPA 432  
|||  
Qy 242 GARRGGASRLPLKRPREGAPEPTPVGGSWAHGRTGPGDRGFCVVSARPA 301  
|||  
Db 433 EEAATSLGALSGRSHSPSVGRQHGAGPPSTSRPPRPWDTPCPVYAETKHFLYSSGDKE 492  
|||  
Qy 302 EEAATSLGALSGRSHSPSVGRQHGAGPPSTSRPPRPWDTPCPVYAETKHFLYSSGDKE 361  
|||  
Db 493 QLRSFLLSLRPSLTGARLVETIFIGSRPMPGTFRRLPRLPQRYWQMRPLFLELLGN 552  
|||  
Qy 362 QLRSFLLSLRPSLTGARLVETIFIGSRPMPGTFRRLPRLPQRYWQMRPLFLELLGN 421  
|||  
Db 553 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRLLVOLLRQHSPPW 612  
|||  
Qy 422 HAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRLLVOLLRQHSPPW 481  
|||  
Db 613 QVYGFVRACLRLVPPGLWGRSHRNERFLRNTKFIISLGHAKLSLQELTWKMSVRDCAW 672  
|||  
Qy 482 QVYGFVRACLRLVPPGLWGRSHRNERFLRNTKFIISLGHAKLSLQELTWKMSVRDCAW 541  
|||  
Db 673 LRRSPGVCVPAEHRUREEILAKFLHMLMSVYVVELLSFFYVTTFTQKNRFFYRKS 732  
|||  
Qy 542 LRRSPGVCVPAEHRUREEILAKFLHMLMSVYVVELLSFFYVTTFTQKNRFFYRKS 601  
|||  
Db 733 VWSKLQSIGIRHKLRYQLRELSAEVQHRREARPAALLTSRLRFPKPDGLRPIVNDYV 792  
|||  
Qy 602 VWSKLQSIGIRHKLRYQLRELSAEVQHRREARPAALLTSRLRFPKPDGLRPIVNDYV 661  
|||  
Db 793 VGARTFRREKRAELTSRVKALFSLVNYERARRPGLLGASVGLGDDIHRARWTFVLVRA 852  
|||  
Qy 662 VGARTFRREKRAELTSRVKALFSLVNYERARRPGLLGASVGLGDDIHRARWTFVLVRA 721  
|||  
Db 853 QDPPPELYFVKVDVGTAYDITIPQDRLEVIASIIKQNTYCVRRYAVVQKAAGHVKAF 912  
|||  
Qy 722 QDPPPELYFVKVDVGTAYDITIPQDRLEVIASIIKQNTYCVRRYAVVQKAAGHVKAF 781  
|||  
Db 913 KSHVSTUTDLQPYMROFVAHLQETSPLRDAVVVIEQSSSLEASSGLDFVLRPMCHAVR 972  
|||

QY 782 KSHVSTUTDLQPYMROFVAHLQETSPLRDAVVVIEQSSSLEASSGLDFVLRPMCHAVR 841  
|||  
Db 973 IRGKSYVQCGIPQGSILSTLLCSLCYGDWENKLFAGIRRDGLLLRLVDLFLVTPHLTH 1032  
|||  
QY 842 IRGKSYVQCGIPQGSILSTLLCSLCYGDWENKLFAGIRRDGLLLRLVDLFLVTPHLTH 901  
|||  
Db 1033 AKTFELTLVRGPEYGCYVNLKRTVVNFVVEALGGTAFVQMPAHGLFPWCGLLDTRT 1092  
|||  
QY 902 AKTFELTLVRGPEYGCYVNLKRTVVNFVVEALGGTAFVQMPAHGLFPWCGLLDTRT 961  
|||  
Db 1093 LEVQSDYSYARTSIRASLTFRNGFKAGNMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1152  
|||  
QY 962 LEVQSDYSYARTSIRASLTFRNGFKAGNMRKLFGLVLRKCHSLFLDLQVNSLQTVCT 1021  
|||  
Db 1153 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVISTASLCYSILKAKNAGMSLGAK 1212  
|||  
QY 1022 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVISTASLCYSILKAKNAGMSLGAK 1081  
|||  
Db 1213 GAAGPLPSEAVOWLCHQAFLLKLTLRHRTVYVPLGLSLRTAQOLSKRLPGTTLTALAAAA 1272  
|||  
QY 1082 GAAGPLPSEAVOWLCHQAFLLKLTLRHRTVYVPLGLSLRTAQOLSKRLPGTTLTALAAAA 1141  
|||  
Db 1273 NPALPSDFKTILD 1285  
|||  
QY 1142 NPALPSDFKTILD 1154  
|||

RESULT 11  
ID US-08-911-312-55 STANDARD; PRT; 1407 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 55, Application US/08911312  
XX  
XX Sequence 55, Application US/08911312  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911,312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997





CC APPLICATION NUMBER: US/08/974,549  
CC FILING DATE: 19-NOV-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 08-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph Ted  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002610US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 628:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1407 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: Protein  
CC LOCATION: 1..1407  
CC OTHER INFORMATION: /note= "fusion protein composed of  
CC OTHER INFORMATION: enhanced green fluorescent protein  
CC OTHER INFORMATION:  
CC OTHER INFORMATION: 5' untranslated region of hTERT mRNA and  
CC OTHER INFORMATION: hTERT protein sequence"  
CC SEQUENCE 1407 AA; 157668 MW; 10134798 CN;

Query Match 99.6%; Score 8590; DB 18; Length 1407;  
Best Local Similarity 99.7%; Pred. No. 0.00e+00;  
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLRTWEALAPATAMPAPRCRAVRSLLRSHYREVLPATFVRRLPGQGWRLV 314  
QY 2 ASGORCVLLRTWEALAPATAMPAPRCRAVRSLLRSHYREVLPATFVRRLPGQGWRLV 61  
Db 315 QRGDPAAFRALVAQCLVCVWDARPPAAPSFROVSCCLKELVARVLQRCERGAKNVLA 374  
QY 62 QRGDPAAFRALVAQCLVCVWDARPPAAPSFROVSCCLKELVARVLQRCERGAKNVLA 121  
Db 375 GFALLDGARGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLVHLLARCALF 434  
QY 122 GFALLDGARGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVVLVHLLARCALF 181

Db 435 VLVAPSCAYOVCGPPLYQLGAATQARPPPHASGPRRRRLGCERAWNHSVREAGVPLGLPAP 494  
QY 182 VLVAPSCAYOVCGPPLYQLGAATQARPPPHASGPRRRRLGCERAWNHSVREAGVPLGLPAP 241  
Db 495 GARRRGGSASRSPLPKRPRRGAAPERTPVGQSWAHPTGRGSDRGFCVVSAPARA 554  
QY 242 GARRRGGSASRSPLPKRPRRGAAPERTPVGQSWAHPTGRGSDRGFCVVSAPARA 301  
Db 555 EBATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPMDTPCPVYVAETKHFLYSSGDKE 614  
QY 302 EBATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPMDTPCPVYVAETKHFLYSSGDKE 361  
Db 615 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGN 674  
QY 362 QLRPSFLLSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGN 421  
Db 675 HAQCPYGVLLKTHCPLRAAATPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRHSSPW 734  
QY 422 HAQCPYGVLLKTHCPLRAAATPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRHSSPW 481  
Db 735 QYGVFVRACLRLVPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 794  
QY 482 QYGVFVRACLRLVPPGLWGSRHNERFLRNTKFTSLGKHAKLSLQELTWKMSVRDCAW 541  
Db 795 LRSPGVCVPAAEHRLREEILAKELHMLMSVYVVELLSFFYVTTTQKRLFFYRPS 854  
QY 542 LRSPGVCVPAAEHRLREEILAKELHMLMSVYVVELLSFFYVTTTQKRLFFYRPS 601  
Db 855 VMSKQSIGIRQHLKRVQLRELSEAEVRQREARPAALLSRRLFIPKPDGLRPIVNDYV 914  
QY 602 VMSKQSIGIRQHLKRVQLRELSEAEVRQREARPAALLSRRLFIPKPDGLRPIVNDYV 661  
Db 915 VGARTFRREKRAERLTSRVKALFSVLNLERARRPGLLGASVLGDDDIHRAWRTFVLVRVA 974  
QY 662 VGARTFRREKRAERLTSRVKALFSVLNLERARRPGLLGASVLGDDDIHRAWRTFVLVRVA 721  
Db 975 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIIPQNTYCVRRYAVVQAAHGHVRKAF 1034  
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIASIIIPQNTYCVRRYAVVQAAHGHVRKAF 781  
Db 1035 KSHVSTLTLQPYMROFVAHLQETSPLRDVAVIEQSSSLEASSGLFDVFLRFMCHHAVR 1094  
QY 782 KSHVSTLTLQPYMROFVAHLQETSPLRDVAVIEQSSSLEASSGLFDVFLRFMCHHAVR 841  
Db 1095 IRKSYVQCGIPQGSILSTLCSLCYGDMEKNKLFAGIRRDGLLLRLYDDFLLVTPHLTH 1154  
QY 842 IRKSYVQCGIPQGSILSTLCSLCYGDMEKNKLFAGIRRDGLLLRLYDDFLLVTPHLTH 901  
Db 1155 AKTFLETLVRGVPYCCVYNLRKTVVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 1214  
QY 902 AKTFLETLVRGVPYCCVYNLRKTVVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 961  
Db 1215 LEVQSDYSSYARTSIRASVTFNRGFKAGRMRRKLFGLVLRKCHSLFLDLQVNSLQTCT 1274  
QY 962 LEVQSDYSSYARTSIRASVTFNRGFKAGRMRRKLFGLVLRKCHSLFLDLQVNSLQTCT 1021  
Db 1275 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVYISDTASLCYSILKAKNAGMSLGAK 1334  
QY 1022 NIYKILLQAYRFHACVQLQPFHQQVKNPTFFLRVYISDTASLCYSILKAKNAGMSLGAK 1081  
Db 1335 GAAGPLPSEAVQWLCHQAFLLKLTLRHRTYVPLLSRLTAQTLQSLRKLPGTTLTALEAAA 1394  
QY 1082 GAAGPLPSEAVQWLCHQAFLLKLTLRHRTYVPLLSRLTAQTLQSLRKLPGTTLTALEAAA 1141  
Db 1395 NPALPSDFKTILD 1407  
QY 1142 NPALPSDFKTILD 1154

RESULT 13  
ID US-08-912-951-334 STANDARD; PRT; 1407 AA.  
XX  
AC xxxxxx



Sequence 334, Application US/08912951  
Sequence 334, Application US/08912951  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
TITLE OF INVENTION: THERAPEUTIC METHODS  
NUMBER OF SEQUENCES: 335  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 334:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1407 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 1407 AA; 157668 MW; 10134798 CN;  
Query Match 99.6%; Score 8590; DB 13; Length 1407;  
Best Local Similarity 99.7%; Pred. No. 0.00e+00;  
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLRTWEALAPATPAMPAPRCRAVRSILRSHYREVLPATFVRLPGQGRVLY 314  
QY 2 ASGORCVLLRTWEALAPATPAMPAPRCRAVRSILRSHYREVLPATFVRLPGQGRVLY 61  
Db 315 QRGDPAAFRALVAQCLVCPVWDARPPAAPSFQVSCLELVARVLRQRCERAKNVLA 374  
QY 62 QRGDPAAFRALVAQCLVCPVWDARPPAAPSFQVSCLELVARVLRQRCERAKNVLA 121  
Db 375 GFALLDGGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLHLLARCALF 434  
QY 122 GFALLDGGARGGPEAFTTSVRSYLPNTVTDALRGSGAWGLLLRRVGDVLLHLLARCALF 181  
Db 435 VLVAPOCAYOVCGPPLYOLGAATQARPPHASPGRRLGRCERANNHVRAGVPLGLPAP 494  
QY 182 VLVAPOCAYOVCGPPLYOLGAATQARPPHASPGRRLGRCERANNHVRAGVPLGLPAP 241  
Db 495 GARRRGGSASRSLPLPKRPRRGAAPERTPVQGGWAHGPGRTRGSDRGFCVVSAPRA 554  
QY 242 GARRRGGSASRSLPLPKRPRRGAAPERTPVQGGWAHGPGRTRGSDRGFCVVSAPRA 301  
Db 555 EATSLGALSGRTHSHPSVGRQHAGPPTSRPPRPWDTPCPVVAETKHFLYSSGDKE 614  
QY 302 EATSLGALSGRTHSHPSVGRQHAGPPTSRPPRPWDTPCPVVAETKHFLYSSGDKE 361  
Db 615 QLPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGN 674  
QY 362 QLPSFLLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLELLGN 421  
Db 675 HAOCPIGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVNAPEEEDTPRRLVOLLRQHSPPW 734  
QY 422 HAOCPIGVLLKTHCPLRAAVTPAAGVCAREKPOGSAVNAPEEEDTPRRLVOLLRQHSPPW 481  
Db 735 QVYGFVRACLRLVPPGLWGRHNERFLRNTKFTISLGHAKLSLQELTKWMSVRDCAW 794  
QY 482 QVYGFVRACLRLVPPGLWGRHNERFLRNTKFTISLGHAKLSLQELTKWMSVRDCAW 541  
Db 795 LRSPGVGCVPAAEHRLREELAKFLHLMVSVVVELLSRFFVYTTTFOKNRLFYFRPS 854  
QY 542 LRSPGVGCVPAAEHRLREELAKFLHLMVSVVVELLSRFFVYTTTFOKNRLFYFRPS 601  
Db 855 VSKLQSIGIRQHLKRVOLRELSEAEVQRHREARPAALLTSRLRFPKPDGLRPIVNMVYV 914  
QY 602 VSKLQSIGIRQHLKRVOLRELSEAEVQRHREARPAALLTSRLRFPKPDGLRPIVNMVYV 661  
Db 915 VGARTFRERKRAEELTSRVKALFVSLNYERARRPGLLGASVLGDDDIHRAWRIFVLVRA 974  
QY 662 VGARTFRERKRAEELTSRVKALFVSLNYERARRPGLLGASVLGDDDIHRAWRIFVLVRA 721  
Db 975 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIAIIPONTYCVRRYAVVQAAHGHVRKAF 1034  
QY 722 QDPPPELYFVKVDVTGAYDTIPQDRLTEVIAIIPONTYCVRRYAVVQAAHGHVRKAF 781  
Db 1035 KSHVSTLTDLPQYMRQFVAHLQETSPURDAVIEQSSSLNEASSGLDFVLFREMFCHAVR 1094  
QY 782 KSHVSTLTDLPQYMRQFVAHLQETSPURDAVIEQSSSLNEASSGLDFVLFREMFCHAVR 841  
Db 1095 IRGKSYVQCGIPIQGSITSLTLLSCYGDMEKLFAGIRDRGLLLRLVDDFLVTPHLTH 1154  
QY 842 IRGKSYVQCGIPIQGSITSLTLLSCYGDMEKLFAGIRDRGLLLRLVDDFLVTPHLTH 901  
Db 1155 AKTFRLTVRGVPEYGCVMNLKRTVNVNFVEDEALGTAFAVQPAHGLFPWCGLLDTRT 1214  
QY 902 AKTFRLTVRGVPEYGCVMNLKRTVNVNFVEDEALGTAFAVQPAHGLFPWCGLLDTRT 961  
Db 1215 LEVQSDYSYARTSIRASVTFNRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQVCT 1274  
QY 962 LEVQSDYSYARTSIRASVTFNRGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQVCT 1021  
Db 1275 NIKYKILLQAYRHFACVLOLPFFHQVWKNPTFFLRVTSDFASLCYSTLKAKNAGMSLGAK 1334  
QY 1022 NIKYKILLQAYRHFACVLOLPFFHQVWKNPTFFLRVTSDFASLCYSTLKAKNAGMSLGAK 1081  
Db 1335 GAAGPLPSEAVQWMLCHQAFLLKLTURHRTVYVPLLGSLRTAQTLRSKRLPGTTLTLEAAA 1394

QY 1082 GAAGPLPSEAWQLCHQAFLLKLFHRYVYVPLGLSLTATQLSRKLPGTTTALAEAA 1141  
Db 1395 NPALPSDFKTLTD 1407  
QY 1142 NPALPSDFKTLTD 1154  
RESULT 14  
ID US-08-974-549-2 STANDARD; PRT; 1132 AA.  
XX xxxxxx  
DE  
DT  
XX Sequence 2, Application US/08974549  
XX Sequence 2, Application US/08974549  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
CC NUMBER OF SEQUENCES: 726  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/974,549  
CC FILING DATE: 19-NOV-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885

CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph Ted  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-00026100S  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1132 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1132 AA; 126995 MW; 6588209 CN;  
Query Match 98.2%; Score 8465; DB 18; Length 1132;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRLRPGQWRLVQSGDPAAFRALVAOCLVCPW 60  
QY 23 MPRAPRCRAVRSLLRSHYREVLPATFVRLRPGQWRLVQSGDPAAFRALVAOCLVCPW 82  
Db 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 120  
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFGFALLDARGGPPPEAFTTSVR 142  
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLVQLGA 180  
QY 143 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLVQLGA 202  
Db 181 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSGASRSPLPKRPRR 240  
QY 203 ATQARPPPHASGPRRLGGERAWNHSVREAGVPLGLPAPGARRRGSGASRSPLPKRPRR 262  
Db 241 GAAPERTPVGQSWAHFGRTRGSDRGFCVVSVPARPAEATSLGALSGRHSPSVG 300  
QY 263 GAAPERTPVGQSWAHFGRTRGSDRGFCVVSVPARPAEATSLGALSGRHSPSVG 322  
Db 301 ROHAGPPSTSRPPRWDTPCPVYATKHFYSSGDKQLRPSFLLSLRPSLTGARRL 360  
QY 323 ROHAGPPSTSRPPRWDTPCPVYATKHFYSSGDKQLRPSFLLSLRPSLTGARRL 382  
Db 361 VETIFLGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGHAQCPYGVLLKTHCPRAAVT 420  
QY 383 VETIFLGSRPMPGTPRRLPRLPQRYWQMRPLFLELLGHAQCPYGVLLKTHCPRAAVT 442  
Db 421 PAAGVCAREKPGQSVAAPEEDTDPRLLVOLLRQHSPPQVYGFVACLRRLVPPGLWGS 480  
QY 443 PAAGVCAREKPGQSVAAPEEDTDPRLLVOLLRQHSPPQVYGFVACLRRLVPPGLWGS 502  
Db 481 RHNERFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHLREEI 540  
QY 503 RHNERFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHLREEI 562  
Db 541 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRLLFFYKRSVWSKLSQIGIROHLKRVQLRE 600  
QY 563 LAKFLHLMSSVYVVELLSRFFYVTTTFOKNRLLFFYKRSVWSKLSQIGIROHLKRVQLRE 622  
Db 601 LSAEVRQHREARPAALLTSKLRFIPKPDGLRPIVNDYVVGARTFREKRAERLTSRKA 660  
QY 623 LSAEVRQHREARPAALLTSKLRFIPKPDGLRPIVNDYVVGARTFREKRAERLTSRKA 682  
Db 661 LFSVLNVERARRPGLLGASVGLGDDIHRARWTFVLVRAQDPPPELYFVKVDVTGAYDTI 720  
QY 683 LFSVLNVERARRPGLLGASVGLGDDIHRARWTFVLVRAQDPPPELYFVKVDVTGAYDTI 742  
Db 721 PODRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVKAFKSHVSTLTDLQPYMRQFVAHL 780  
QY 743 PODRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVKAFKSHVSTLTDLQPYMRQFVAHL 802  
Db 781 QETSPLEDAVVEIQSSSLNEASSGLFDVFLRFMCHHAVIRNGKSYVQCQGIPOGSIILSTL 840

QY 803 QETSPRLDAWVIEQSSSLNEASSGLEDVFLFENCHAVIRKSYVQCOGIPGSGILSTL 862  
DB 841 LCSLCYGDMEKLFAGIRRDGLLLRLVDDFLVTPHLTHAKTFLRLVRGVPEYGCWNL 900  
QY 863 LCSLCYGDMEKLFAGIRRDGLLLRLVDDFLVTPHLTHAKTFLRLVRGVPEYGCWNL 922  
DB 901 RKTUVNFPVEDEALGCTAFQVMPAHGLFPWCGLLDTRILEVQSDYSYARTSIRASLTF 960  
QY 923 RKTUVNFPVEDEALGCTAFQVMPAHGLFPWCGLLDTRILEVQSDYSYARTSIRASLTF 982  
DB 961 NRGFKAGRNRRKLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQAYRHACVQLP 1020  
QY 983 NRGFKAGRNRRKLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQAYRHACVQLP 1042  
DB 1021 FHOQVWKNPTFFLURVSDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLC HQAFLL 1080  
QY 1043 FHOQVWKNPTFFLURVSDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLC HQAFLL 1102  
DB 1081 KLTRHRTVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAANPALPSDFKTILD 1132  
QY 1103 KLTRHRTVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAANPALPSDFKTILD 1154

RESULT 15  
ID US-08-912-951-2 STANDARD; PRT; 1132 AA.  
XX xxxxxx  
Sequence 2, Application US/08912951  
Sequence 2, Application US/08912951  
GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-0026000US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1132 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1132 AA; 126995 MW; 6588209 CN;  
Query Match 98.2%; Score 8465; DB 13; Length 1132;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLVQRGDPAAFRALVAQCILVCVPW 60  
QY 23 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLVQRGDPAAFRALVAQCILVCVPW 82  
DB 61 DARPPPAAPSFQVSVCLKELVARVQLRCERGAKNVLAFFGALLDARGGPPPEAFTTSVR 120  
QY 83 DARPPPAAPSFQVSVCLKELVARVQLRCERGAKNVLAFFGALLDARGGPPPEAFTTSVR 142  
DB 121 SYLPTNTVDALRGSGAWGLLRRVGGDVLVLLARCALFVLVAPSCAYQVCGPPPLYQLGA 180  
QY 143 SYLPTNTVDALRGSGAWGLLRRVGGDVLVLLARCALFVLVAPSCAYQVCGPPPLYQLGA 202  
DB 181 ATQARPPPHASGPRRRRLGGERAWNSVREAGVPLGLPAPGARRRGGSASRSPLPKRPRR 240  
QY 203 ATQARPPPHASGPRRRRLGGERAWNSVREAGVPLGLPAPGARRRGGSASRSPLPKRPRR 262  
DB 241 GAAPPERTPVQGSWAHPGTRGPDGFCVVPSPARPAEATSLLEGALSGTRHSHPSVG 300  
QY 263 GAAPPERTPVQGSWAHPGTRGPDGFCVVPSPARPAEATSLLEGALSGTRHSHPSVG 322  
DB 301 RQHAGPPSTSRPPRPWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARRL 360  
QY 323 RQHAGPPSTSRPPRPWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRSLTGARRL 382  
DB 361 VETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFELLGNHACPYGVLLKTHCPLRAAVT 420  
QY 383 VETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFELLGNHACPYGVLLKTHCPLRAAVT 442  
DB 421 PAAGVAREKPGQSVAAPEEDTDPRRLVQLLRQHSQWVYGFVRACLRRLVPPGLMGS 480  
QY 443 PAAGVAREKPGQSVAAPEEDTDPRRLVQLLRQHSQWVYGFVRACLRRLVPPGLMGS 502  
DB 481 RNERREFLNTKKFISLGKHAHLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHLRREI 540  
QY 503 RNERREFLNTKKFISLGKHAHLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHLRREI 562  
DB 541 LAKFLHMLMSVYVVELLSFFVYTTTTFQKNRLLFFYKRSVWSKLSIGIRQHLKRVQLRE 600  
QY 563 LAKFLHMLMSVYVVELLSFFVYTTTTFQKNRLLFFYKRSVWSKLSIGIRQHLKRVQLRE 622  
DB 601 LSEAEVQRHREARPAALLTSRLRFIPKPDGLRFINVMDYVVGARTFRREKRAERLTSRVKA 660  
QY 623 LSEAEVQRHREARPAALLTSRLRFIPKPDGLRFINVMDYVVGARTFRREKRAERLTSRVKA 682  
DB 661 LFSVLNFERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELYFKVVDVTGAYDTI 720

QY 683 LFSVLNVERARRRGLGASVLGDDIHRWRTFVLVRQAQDPPPELYFVKVDVTGAYDI 742  
Db 721 PODRLTEVIASIIKQNTYCVRRYAVVQKAAGHVRKAFKSHYSTLTLDLPYMRQFVAHL 780  
QY 743 PODRLTEVIASIIKQNTYCVRRYAVVQKAAGHVRKAFKSHYSTLTLDLPYMRQFVAHL 802  
Db 781 QETSPDLADAVVIEQSSSLNEASSGLFDFVRFCMHAVRIRGKSYVQCQIPOGSIILSTL 840  
QY 803 QETSPDLADAVVIEQSSSLNEASSGLFDFVRFCMHAVRIRGKSYVQCQIPOGSIILSTL 862  
Db 841 LCSLCYGDMEKILFAGIRRRDGLLRVDDFLLVTPHLTHAKTFLRTLVRGVEYGCVVNL 900  
QY 863 LCSLCYGDMEKILFAGIRRRDGLLRVDDFLLVTPHLTHAKTFLRTLVRGVEYGCVVNL 922  
Db 901 RKTWNFPVEDEALGTAFAVQMPAHGLFPMCGLLDTRTLEVOSSDYSSVARTSIRASLTF 960  
QY 923 RKTWNFPVEDEALGTAFAVQMPAHGLFPMCGLLDTRTLEVOSSDYSSVARTSIRASLTF 982  
Db 961 NRGEKAGNRMRKLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQAYRFHACVLQLP 1020  
QY 983 NRGEKAGNRMRKLFGLVRLKCHSLFELDLQVNSLQVCTNIYKILLQAYRFHACVLQLP 1042  
Db 1021 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWMLCHOAFL 1080  
QY 1043 FHOQWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWMLCHOAFL 1102  
Db 1081 KLTRHRVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAAANPALPSDFKTILD 1132  
QY 1103 KLTRHRVYVPLGLSLRTAQTLQSRKLPGLTTLTALEAAANPALPSDFKTILD 1154

RESULT 16  
ID US-08-911-312-2 STANDARD: PRT: 1132 AA.  
XX  
AC  
XX  
XX  
DT  
XX  
Sequence 2, Application US/08911312  
Sequence 2, Application US/08911312  
GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911,312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-0025000S  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1132 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE: 1132 AA; 126995 MW; 6588209 CN;  
Query Match 98.2%; Score 8465; DB 13; Length 1132;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLLGPQWRVLVQRGDPAAFRALVAQCLVCVPW 60  
QY 23 MPRAPRCRAVRSLLRSHYREVLPATFVRRLLGPQWRVLVQRGDPAAFRALVAQCLVCVPW 82  
Db 61 DARPPPAAPSFQVQSCLELVARVLQRCERGAKNVLAFFGALLDGGARPPPEATTTSVR 120  
QY 83 DARPPPAAPSFQVQSCLELVARVLQRCERGAKNVLAFFGALLDGGARPPPEATTTSVR 142  
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYOVCGPPLYQLGA 180  
QY 143 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYOVCGPPLYQLGA 202  
Db 181 ATQARPPPHASCPRRRLGGERAWNHSVREAGVPLGLPAGARRRGGSASRSLPLKRRPR 240  
QY 203 ATQARPPPHASCPRRRLGGERAWNHSVREAGVPLGLPAGARRRGGSASRSLPLKRRPR 262  
Db 241 GAAPERTPVQGSWAHPGRTGRGSDRGFCVVSAPAEAEATSLGALSGTRHSHPSVG 300  
QY 263 GAAPERTPVQGSWAHPGRTGRGSDRGFCVVSAPAEAEATSLGALSGTRHSHPSVG 322  
Db 301 ROHAGPPSTSRPPRPWDTPCPVYAEKHFYSSGDEQLRPSFLLSRLRSLTGARRL 360  
QY 323 ROHAGPPSTSRPPRPWDTPCPVYAEKHFYSSGDEQLRPSFLLSRLRSLTGARRL 382  
Db 361 VETIFLGSRRPMPTPRRLPLRPLQRYWOMRPLFLELGNHQAOCYPYVLLKTHCPRAAVT 420  
QY 383 VETIFLGSRRPMPTPRRLPLRPLQRYWOMRPLFLELGNHQAOCYPYVLLKTHCPRAAVT 442  
Db 421 PAAGVCAREKPOGSVAAPDEEDTDPRLVQLLRQHSHPQVYGFVRACLRRLRVPPLGWS 480  
QY 443 PAAGVCAREKPOGSVAAPDEEDTDPRLVQLLRQHSHPQVYGFVRACLRRLRVPPLGWS 502  
Db 481 RHNRERFLNRTKKFISLKGKHAQLSLQELTWKMSVRDCAWLRKSPGVGCVPAAEHRLREI 540  
QY 503 RHNRERFLNRTKKFISLKGKHAQLSLQELTWKMSVRDCAWLRKSPGVGCVPAAEHRLREI 562  
Db 541 LAKFLHLMMSVYVVELLSRFFVYTTETQKNELFFYRKSWSKLOSIGIRQHLKRVOLRE 600  
QY 563 LAKFLHLMMSVYVVELLSRFFVYTTETQKNELFFYRKSWSKLOSIGIRQHLKRVOLRE 622

Db	601	L	S	E	A	E	V	R	O	H	R	E	A	R	P	A	L	L	T	S	R	L	F	I	P	K	P	D	G	L	R	I	V	N	M	D	V	V	G	A	R	T	F	R	E	K	R	A	E	R	L	T	S	R	V	K	A	660
QY	623	L	S	E	A	E	V	R	O	H	R	E	A	R	P	A	L	L	T	S	R	L	F	I	P	K	P	D	G	L	R	I	V	N	M	D	V	V	G	A	R	T	F	R	E	K	R	A	E	R	L	T	S	R	V	K	A	682
Db	661	L	F	S	V	L	N	Y	E	R	A	R	P	G	L	L	G	A	S	V	L	G	D	D	I	H	R	A	W	T	F	V	L	R	V	R	A	O	P	P	P	E	L	F	V	K	V	D	T	G	A	Y	D	T	I	720		
QY	683	L	F	S	V	L	N	Y	E	R	A	R	P	G	L	L	G	A	S	V	L	G	D	D	I	H	R	A	W	T	F	V	L	R	V	R	A	O	P	P	P	E	L	F	V	K	V	D	T	G	A	Y	D	T	I	742		
Db	721	P	O	B	R	L	E	V	I	A	S	I	I	K	P	O	N	T	V	C	V	R	Y	A	V	O	K	A	H	G	H	V	R	K	A	F	K	S	H	V	S	T	L	D	L	O	P	Y	M	R	Q	F	A	H	L	780		
QY	743	P	O	B	R	L	E	V	I	A	S	I	I	K	P	O	N	T	V	C	V	R	Y	A	V	O	K	A	H	G	H	V	R	K	A	F	K	S	H	V	S	T	L	D	L	O	P	Y	M	R	Q	F	A	H	L	802		
Db	781	Q	E	T	S	P	L	R	D	A	V	I	E	O	S	S	L	N	E	A	S	G	L	F	D	V	F	L	R	M	C	H	H	A	V	I	R	G	K	S	V	O	C	O	G	I	P	O	G	S	I	L	T	L	840			
QY	803	Q	E	T	S	P	L	R	D	A	V	I	E	O	S	S	L	N	E	A	S	G	L	F	D	V	F	L	R	M	C	H	H	A	V	I	R	G	K	S	V	O	C	O	G	I	P	O	G	S	I	L	T	L	862			
Db	841	L	C	S	I	C	Y	G	D	M	E	N	K	L	F	A	G	I	R	D	G	L	L	L	V	D	D	F	L	V	T	P	H	L	T	H	A	K	T	F	L	T	V	R	G	V	P	E	Y	G	C	V	V	N	L	900		
QY	863	L	C	S	I	C	Y	G	D	M	E	N	K	L	F	A	G	I	R	D	G	L	L	L	V	D	D	F	L	V	T	P	H	L	T	H	A	K	T	F	L	T	V	R	G	V	P	E	Y	G	C	V	V	N	L	922		
Db	901	R	K	T	V	N	F	P	E	D	E	A	L	G	T	A	F	V	O	M	P	A	N	G	L	F	P	W	C	G	L	L	D	T	T	L	E	V	O	S	D	S	Y	S	I	A	R	T	S	I	R	A	S	L	T	F	960	
QY	923	R	K	T	V	N	F	P	E	D	E	A	L	G	T	A	F	V	O	M	P	A	N	G	L	F	P	W	C	G	L	L	D	T	T	L	E	V	O	S	D	S	Y	S	I	A	R	T	S	I	R	A	S	L	T	F	982	
Db	961	N	R	G	F	K	A	G	R	N	M	R	K	L	F	G	V	L	R	K	C	H	S	L	F	L	D	L	O	V	N	S	L	O	T	C	T	N	I	Y	K	L	L	Q	A	R	F	H	A	C	V	L	O	L	P	1020		
QY	983	N	R	G	F	K	A	G	R	N	M	R	K	L	F	G	V	L	R	K	C	H	S	L	F	L	D	L	O	V	N	S	L	O	T	C	T</																					

RESULT 17

ID	US-09-042-460-3	STANDARD;	PRT; 1132 AA.
----	-----------------	-----------	---------------

AC XXXXXX

XX

DT

XX DE

XX  
XX

Sequencer

CC  
GÈNÈRCC  
APCC  
APICC  
APR22  
21CC  
NDU

102

33

33

22

888

333

33

100

33

55

CU 55

33

22

CC  
PR:

22

CC	FILING DATE: 01-OCT-1996
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/844,419
CC	FILING DATE: 18-APR-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/846,017
CC	FILING DATE: 25-APR-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/851,843
CC	FILING DATE: 06-MAY-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/854,050
CC	FILING DATE: 09-MAY-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/911,312
CC	FILING DATE: 14-AUG-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/912,951
CC	FILING DATE: 14-AUG-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/915,503
CC	FILING DATE: 14-AUG-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: WO PCT/US97/17618
CC	FILING DATE: 01-OCT-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: WO PCT/US97/17885
CC	FILING DATE: 01-OCT-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/974,549
CC	FILING DATE: 19-NOV-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/974,584
CC	FILING DATE: 19-NOV-1997
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 08/979,742
CC	FILING DATE: 26-NOV-1997
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Einhorn, Gregory P.
CC	REGISTRATION NUMBER: 38,440
CC	REFERENCE/DOCKET NUMBER: 015389-00311
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (415) 576-0200
CC	TELEFAX: (415) 576-0300
CC	INFORMATION FOR SEQ ID NO: 3:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 1132 amino acids
CC	TYPE: amino acid
CC	STRANDEDNESS:
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: protein
CC	FEATURE:
CC	NAME/KEY: Protein
CC	LOCATION: 1..1132
CC	OTHER INFORMATION: /note= "human telom
CC	OTHER INFORMATION: transcriptase (hTR
CC	SEQUENCE 1132 AA; 126995 MW; 6588209 CN;
CC	SEQUENCE
CC	Query Match 98.2%; Score 8465; DB 1
CC	Best Local Similarity 100.0%; Pred. No. 0.00e+
CC	Matches 1132; Conservative 0; Mismatches
CC	1 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLV
CC	1132
CC	23 MPRAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRVLV
CC	1132
CC	61 DAREPPPAAPSFQVSCSLKELVARVLQRLCERGAKNVLAF
CC	1132
CC	83 DAREPPPAAPSFQVSCSLKELVARVLQRLCERGAKNVLAF
CC	1132
CC	121 SYLPNTVTDALRGSGANGLLLRVGDVVLVHLRLARCAFT
CC	1132
CC	143 SYLPNTVTDALRGSGANGLLLRVGDVVLVHLRLARCAFT
CC	1132

Db 191 ATQARPPPHASGPRRRRLGGERAWNHVSREAGVPLGLPAGARRRGGASRSLPLPKRPRR 240  
Qy 203 ATQARPPPHASGPRRRRLGGERAWNHVSREAGVPLGLPAGARRRGGASRSLPLPKRPRR 262  
Db 241 GAAPERTPVGQSWAHPTGTRGSDGFCVSPARPAEATSELEGALSCTRSHPSVG 300  
Qy 263 GAAPERTPVGQSWAHPTGTRGSDGFCVSPARPAEATSELEGALSCTRSHPSVG 322  
Db 301 ROHAGPSTSRPRPMDTPCPVYAEKHELYSSGDEQRLPFLSSLRPSLTGARRL 360  
Qy 323 ROHAGPSTSRPRPMDTPCPVYAEKHELYSSGDEQRLPFLSSLRPSLTGARRL 382  
Db 361 VETIFLGRPMWPGTRPRRLPRLPQRYWQMRPLFLELLGNHACQPYGVLLKTHCPRLAAVT 420  
Qy 383 VETIFLGRPMWPGTRPRRLPRLPQRYWQMRPLFLELLGNHACQPYGVLLKTHCPRLAAVT 442  
Db 421 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRHSSPWQYGVFRACLRRLVPPGLWGS 480  
Qy 443 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRHSSPWQYGVFRACLRRLVPPGLWGS 502  
Db 481 RHNERFLRNTKFTSLGKHAKLSLOELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 540  
Qy 503 RHNERFLRNTKFTSLGKHAKLSLOELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI 562  
Db 541 LAKELHMLMSVYVELLSRFFVTETTFQKNRLEFFYKRSVMSKLSQSIGIRHQLKRVQLE 600  
Qy 563 LAKELHMLMSVYVELLSRFFVTETTFQKNRLEFFYKRSVMSKLSQSIGIRHQLKRVQLE 622  
Db 601 LSEAEVROHREARPAALLTSRLRFPKPDGLRPIVNMVYVGCARTFRKRAERLTSRVKA 660  
Qy 623 LSEAEVROHREARPAALLTSRLRFPKPDGLRPIVNMVYVGCARTFRKRAERLTSRVKA 682  
Db 661 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELVKVDVTGADTI 720  
Qy 683 LFSVLNTERARRPGLLGASVLGLDDIHRAWRTFVLVRAQDPPPELVKVDVTGADTI 742  
Db 721 PODRLTEVIASIIKPNQTYCVRRYAVVOKAAHGVKRAFKSHVSTLTDLPYMRQFVAHL 780  
Qy 743 PODRLTEVIASIIKPNQTYCVRRYAVVOKAAHGVKRAFKSHVSTLTDLPYMRQFVAHL 802  
Db 781 QETSPLRDAVVEQSSLINEASSGLFDFVLFPMCHHVRIRGKSVOCQIPQGSILSTL 840  
Qy 803 QETSPLRDAVVEQSSLINEASSGLFDFVLFPMCHHVRIRGKSVOCQIPQGSILSTL 862  
Db 841 LCSLCYGDMEKLFAGIRRDGILLRLVDDFLVTPHLLTHAKTFLRTLVRGYPEYGVVNL 900  
Qy 863 LCSLCYGDMEKLFAGIRRDGILLRLVDDFLVTPHLLTHAKTFLRTLVRGYPEYGVVNL 922  
Db 901 RKTVVNPFVEDEALGGTAFVQMPAHGLFPWCGLLDDTLRTLEVQSDYSYARTSTRASLTF 960  
Qy 923 RKTVVNPFVEDEALGGTAFVQMPAHGLFPWCGLLDDTLRTLEVQSDYSYARTSTRASLTF 982  
Db 961 NRGEKAGNMRKLFGLVLRKCHSLFDLVQNSLOTCTNYKLLILOAYRFHACVQLP 1020  
Qy 983 NRGEKAGNMRKLFGLVLRKCHSLFDLVQNSLOTCTNYKLLILOAYRFHACVQLP 1042  
Db 1021 FHQWKNPTFFLRVISTASICYSLKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1080  
Qy 1043 FHQWKNPTFFLRVISTASICYSLKAKNAGMSLGAAGPLPSEAVOWLCHQAFLL 1102  
Db 1081 KLTHRVTVVPLGLSLRTAQQLSRKPLGPTTLTALEAAANPALPSDFKTIIL 1132  
Qy 1103 KLTHRVTVVPLGLSLRTAQQLSRKPLGPTTLTALEAAANPALPSDFKTIIL 1154

RESULT 18  
ID US-09-052-919-2 STANDARD: PRT: 1132 AA.

AC xxxxxx

XX

DT

XX.

DE Sequence 2, Application US/09052919  
XX Sequence 2, Application US/09052919  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Antisense Compositions for Detecting and  
CC NUMBER OF SEQUENCES: 72  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/052,919  
CC FILING DATE: 31-MAR-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,549  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,584  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Parent, Annette S.  
CC REGISTRATION NUMBER: 42,058  
CC REFERENCE/DOCKET NUMBER: 015389-0036000US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1132 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 1132 AA; 126995 MW; 6588209 CN;

Query Match

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	1	MPRAPCRVRSLLRSHYREVLPVPLATEVRLGQGWLVORGDPAARFALVAQCLVCPW	60
QY	23	MPRAPCRVRSLLRSHYREVLPVPLATEVRLGQGWLVORGDPAARFALVAQCLVCPW	82
Db	61	DARPPAAPSFQVSCVSKELVARVQLRCERGAKNVLAFAFALLDARGGPPPEAFTTSVR	120
QY	83	DARPPAAPSFQVSCVSKELVARVQLRCERGAKNVLAFAFALLDARGGPPPEAFTTSVR	142
Db	121	SYLPNTVTDALRGSGWGLLLRRVDDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA	180
QY	143	SYLPNTVTDALRGSGWGLLLRRVDDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA	202
Db	181	ATQARPPHAGSRPRRLGGERANNHVSREAGVPLGIPAPGARRRGGASRSPLPKRPRR	240
QY	203	ATQARPPHAGSRPRRLGGERANNHVSREAGVPLGIPAPGARRRGGASRSPLPKRPRR	262
Db	241	GAPEPERTVPGGSAWHPGTRGSPDRGFCVVSPPARPAEATSEALSGTRHSHPSVG	300
QY	263	GAPEPERTVPGGSAWHPGTRGSPDRGFCVVSPPARPAEATSEALSGTRHSHPSVG	322
Db	301	RQHAGPSTSRPPRPWDTPCPVPVYAEKHFYSSGDKQLRPSFLSSLRPSLTGARRL	360
QY	323	RQHAGPSTSRPPRPWDTPCPVPVYAEKHFYSSGDKQLRPSFLSSLRPSLTGARRL	382
Db	361	VERIFLGSRPWPGTPRRLPRLPQRYQWMPLELLELGNHAQCPYGVLLTHCPRAAVT	420
QY	383	VERIFLGSRPWPGTPRRLPRLPQRYQWMPLELLELGNHAQCPYGVLLTHCPRAAVT	442
Db	421	PAAGVCAREKPOGSVAAPPEEDTPRLVOLLROHSSPMQVGFVRACLRRLLVPPGLWGS	480
QY	443	PAAGVCAREKPOGSVAAPPEEDTPRLVOLLROHSSPMQVGFVRACLRRLLVPPGLWGS	502
Db	481	RHNERFLRNTKFIISLGHAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI	540
QY	503	RHNERFLRNTKFIISLGHAKLSLQELTWKMSVRDCAWLRSPGVCVPAAEHRLREEI	562
Db	541	LAKFLHLMSSVYVVELLSRFFYVTTETFOKNRLFYRKSWKLSQSGIRQHLKRVOLRE	600
QY	563	LAKFLHLMSSVYVVELLSRFFYVTTETFOKNRLFYRKSWKLSQSGIRQHLKRVOLRE	622
Db	601	LSAEVROHREARPAALLTSRLRFTPKDGLRPVNMDDYVVGARTFREKRAERLTSRVKA	660
QY	623	LSAEVROHREARPAALLTSRLRFTPKDGLRPVNMDDYVVGARTFREKRAERLTSRVKA	682
Db	661	LFSVLNTERARRPGLLGASVLGDDIHRANWTFVLRVRAQDPPPELVFVKVDVTGAYDTI	720
QY	683	LFSVLNTERARRPGLLGASVLGDDIHRANWTFVLRVRAQDPPPELVFVKVDVTGAYDTI	742
Db	721	PQRLTEVIASIKPQNTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL	780
QY	743	PQRLTEVIASIKPQNTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL	802
Db	781	QETSPLRDADVIEQSSSLNEASSGLFDVLFREWMCHHAVIRGKSYVOCQIPOSILSTL	840
QY	803	QETSPLRDADVIEQSSSLNEASSGLFDVLFREWMCHHAVIRGKSYVOCQIPOSILSTL	862
Db	841	LCSLCYGDMENKLFAGIRRDGGLLLRLVDDLLVTPHLLTHAKTFLRTLVRGVPYEGCVVNL	900
QY	863	LCSLCYGDMENKLFAGIRRDGGLLLRLVDDLLVTPHLLTHAKTFLRTLVRGVPYEGCVVNL	922
Db	901	RKTVVNFPEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVDSDYSYARTSIRASLTF	960

QY	923	RKTVVNFPEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVDSDYSYARTSIRASLTF	982
Db	961	NRGFKAGRNRRRLFGVLRKCHSLFELDQVNSLQVCTVCTNIYKILLIQAYRFHACVQLQP	1020
QY	983	NRGFKAGRNRRRLFGVLRKCHSLFELDQVNSLQVCTVCTNIYKILLIQAYRFHACVQLQP	1042
Db	1021	FHOQVKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL	1080
QY	1043	FHOQVKNPTFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHQAFLL	1102
Db	1081	KLFRHRVTVVPLGLSLRTAQOTLSRKLPGTTLTALAAAANPALPSDFKTILD	1132
QY	1103	KLFRHRVTVVPLGLSLRTAQOTLSRKLPGTTLTALAAAANPALPSDFKTILD	1154
RESULT	19		
ID	US-08-854-050-225	STANDARD;	PRT; 1132 AA.
XX	xxxxxx		
DT			
XX			
DE			
XX			
CC	Sequence 225, Application US/08854050		
CC	Sequence 225, Application US/08854050		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Cech, Thomas R.		
CC	APPLICANT: Lingner, Joachim		
CC	APPLICANT: Nakamura, Toru		
CC	APPLICANT: Chapman, Karen B.		
CC	APPLICANT: Morin, Gregg B.		
CC	APPLICANT: Andrews, William H.		
CC	TITLE OF INVENTION: Novel Telomerase		
CC	NUMBER OF SEQUENCES: 225		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESS: Townsend and Townsend and Crew LLP		
CC	STREET: Two Embarcadero Center, 8th Floor		
CC	CITY: San Francisco		
CC	STATE: California		
CC	COUNTRY: United States of America		
CC	ZIP: 94111		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/854,050		
CC	FILING DATE: 09-MAY-1997		
CC	CLASSIFICATION: 536		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/851,843		
CC	FILING DATE: 06-MAY-1997		
CC	CLASSIFICATION: 536		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/846,017		
CC	FILING DATE: 25-APR-1997		
CC	CLASSIFICATION: 536		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/844,419		
CC	FILING DATE: 18-APR-1997		
CC	CLASSIFICATION: 536		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/724,643		
CC	FILING DATE: 01-OCT-1996		
CC	CLASSIFICATION: 536		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Apple, Randolph T.		
CC	REGISTRATION NUMBER: 36,429		
CC	REFERENCE/DOCKET NUMBER: 015389-0029300S		
CC	TELECOMMUNICATION INFORMATION:		



CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 225:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1132 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1132 AA; 126995 MW; 6588209 CN;

Query Match 98.2%; Score 8465; DB 12; Length 1132;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPRAPRCRAVRLSLRSHYREVLPVPLATFVRRLGPOGWRLVORGDPAAFRALVAOCLVCPW 60  
QY 23 MPRAPRCRAVRLSLRSHYREVLPVPLATFVRRLGPOGWRLVORGDPAAFRALVAOCLVCPW 82

Db 61 DARPPPAAPSPFROYSCUKELVARVLQRLCERGAKNVLAFFALLDARGGPPPEAFTTSVR 120  
QY 83 DARPPPAAPSPFROYSCUKELVARVLQRLCERGAKNVLAFFALLDARGGPPPEAFTTSVR 142

Db 121 SYLPTNTDALRGSGANGLLRRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYOLGA 180  
QY 143 SYLPTNTDALRGSGANGLLRRVDDVLLHLLARCALFVLVAPSCAYQVCGPPLYOLGA 202

Db 181 ATQARPPPHASGPRRLRGCRANMHSVREAGVPLGLPAPGARRRGSGASRLPLPKRPRR 240  
QY 203 ATQARPPPHASGPRRLRGCRANMHSVREAGVPLGLPAPGARRRGSGASRLPLPKRPRR 262

Db 241 GAPEPERTVPGSGWAHPGRTGPGDRGFCVVSPPARPAEATSEALSGSTRHSHPSVG 300  
QY 263 GAPEPERTVPGSGWAHPGRTGPGDRGFCVVSPPARPAEATSEALSGSTRHSHPSVG 322

Db 301 RQHAGPSTSRPPRPWDTPCPVYATKHFLLSSGKEQLRPSFLSSLRPSLTGARRL 360  
QY 323 RQHAGPSTSRPPRPWDTPCPVYATKHFLLSSGKEQLRPSFLSSLRPSLTGARRL 382

Db 361 VETIFLSRPMPTGPRRLPLPQRYWOMRPLFLELLGNHAQCPYGYLLKTHCPRAAVT 420  
QY 383 VETIFLSRPMPTGPRRLPLPQRYWOMRPLFLELLGNHAQCPYGYLLKTHCPRAAVT 442

Db 421 PAAGVCAREKPOGSAVAPEEDTPRRLVOLLROHSSPMQVGFVRACTRLRYPPLGWS 480  
QY 443 PAAGVCAREKPOGSAVAPEEDTPRRLVOLLROHSSPMQVGFVRACTRLRYPPLGWS 502

Db 481 RHNERFLRNTKRFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 540  
QY 503 RHNERFLRNTKRFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 562

Db 541 LAKFLHLMVSVYVVELLRSFFYVTTFTQNRLLFFYRKSVMSKLSQSTIGIROHLKRVOLRE 600  
QY 563 LAKFLHLMVSVYVVELLRSFFYVTTFTQNRLLFFYRKSVMSKLSQSTIGIROHLKRVOLRE 622

Db 601 LSEAEVQHEARPAALTSRLRFTPKDGLRPVNMVYVVGARTFREKRAERLTSRVKA 660  
QY 623 LSEAEVQHEARPAALTSRLRFTPKDGLRPVNMVYVVGARTFREKRAERLTSRVKA 682

Db 661 LFSVLNTERARRPCLLGASVLGLDDIHRAWRTFVLVRAODPPPELVFVKVDVTGAYDTI 720  
QY 683 LFSVLNTERARRPCLLGASVLGLDDIHRAWRTFVLVRAODPPPELVFVKVDVTGAYDTI 742

Db 721 PQDRLTEVIAIIPQNTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHL 780  
QY 743 PQDRLTEVIAIIPQNTYCVRRYAVVQKAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHL 802

Db 781 QETSPLRDVAVIEOSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIPQGSILSTL 840  
QY 803 QETSPLRDVAVIEOSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIPQGSILSTL 862

Db 841 LCSLCYGDMEKNKIFAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPEYGCVVNL 900  
QY 863 LCSLCYGDMEKNKIFAGIRRRDGLLRLVDDFLVTPHLLTHAKTFLRTLVRGVPEYGCVVNL 922

Db 901 RKTVVNFPVEDEALGCTAFVQMPAHGLFPWCGLLLDTRILEVOSDYSSYARTSIRASLTF 960  
QY 923 RKTVVNFPVEDEALGCTAFVQMPAHGLFPWCGLLLDTRILEVOSDYSSYARTSIRASLTF 982

Db 961 NRGFKAGRNMRRLFGVLRKCHSLFSLDQVNSLOTVCTNIYKILLQAYRFHACVQLP 1020  
QY 983 NRGFKAGRNMRRLFGVLRKCHSLFSLDQVNSLOTVCTNIYKILLQAYRFHACVQLP 1042

Db 1021 FHQOVKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWHLCHOAFL 1080  
QY 1043 FHQOVKNPTFFFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWHLCHOAFL 1102

Db 1081 KLTRHRVTVPLGLSRTAQTLRSKLPGLTTLTALEAAANPALPSDFKTILD 1132  
QY 1103 KLTRHRVTVPLGLSRTAQTLRSKLPGLTTLTALEAAANPALPSDFKTILD 1154

RESULT 20  
ID US-08-912-951-323 STANDARD; PRT; 1154 AA.  
XX  
AC xxxxxx  
DT  
XX  
XX  
DE Sequence 323, Application US/08912951  
XX  
CC Sequence 323, Application US/08912951  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITLE OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC PRIOR APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC PRIOR APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC PRIOR APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC PRIOR APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC PRIOR APPLICATION NUMBER: US 08/724,643



CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 323:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1154 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1154 AA; 129603 MW; 6816882 CN;

Query Match 98.2%; Score 8465; DB 13; Length 1154;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAQCIVCPW 60  
QY 23 MPAPRCRAVRSLLRSHYREVLPATFVRRLGPGQWRLVQRGDPAAFRALVAQCIVCPW 82

Db 61 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFCFALLDGCARGGPPPEAFTTSVR 120  
QY 83 DARPPPAAPSFQVSCLELVARVQLRCERGAKNVLAFCFALLDGCARGGPPPEAFTTSVR 142

Db 121 SYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180  
QY 143 SYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202

Db 181 ATQARPPPHASGPRRLGGERAWNVSRVAGVPLGLPAPGARRRGGSASRLPLPKRPRR 240  
QY 203 ATQARPPPHASGPRRLGGERAWNVSRVAGVPLGLPAPGARRRGGSASRLPLPKRPRR 262

Db 241 GAAPEPTPVGGSWAHFPGRTGSDRGFCVVSPARPAEATSLGALSGTRHSHPSVG 300  
QY 263 GAAPEPTPVGGSWAHFPGRTGSDRGFCVVSPARPAEATSLGALSGTRHSHPSVG 322

Db 301 ROHHAGPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRPSLTGARLL 360  
QY 323 ROHHAGPSTSRPPRWDTPCPVYAEKHFYSSGDKQLRPSFLLSLRPSLTGARLL 382

Db 361 VETIFLGSPPWMPGTPRRLPLPQRYWQMRPLFLELLGHNACPYGVLLKTHCPLRAAVT 420  
QY 383 VETIFLGSPPWMPGTPRRLPLPQRYWQMRPLFLELLGHNACPYGVLLKTHCPLRAAVT 442

Db 421 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPQVYGFVRACLRRLVPPGLWGS 480  
QY 443 PAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPQVYGFVRACLRRLVPPGLWGS 502

Db 481 RHNERFLNTKFFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 540  
QY 503 RHNERFLNTKFFISLGHAKLSLQELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 562

Db 541 LAKFLHLMASVYVVELLSFFVYTTTFQKNRLFYFKYSWWSKLSQIGIRQHLKRVQLRE 600  
QY 563 LAKFLHLMASVYVVELLSFFVYTTTFQKNRLFYFKYSWWSKLSQIGIRQHLKRVQLRE 622

Db 601 LSEAEVQHQREARPAALLTSRLRFPKPDGLRPIVNDYVVGARTFRREKRAERLTSRKA 660  
QY 623 LSEAEVQHQREARPAALLTSRLRFPKPDGLRPIVNDYVVGARTFRREKRAERLTSRKA 682

Db 661 LFSVLNFERARRPGLLGASVLGLDDTHRAWRFTVLVRAQDPPPELYFKVVDVTGAYDTI 720  
QY 683 LFSVLNFERARRPGLLGASVLGLDDTHRAWRFTVLVRAQDPPPELYFKVVDVTGAYDTI 742

Db 721 PDRLTEVIASIIKPPONTYCVRYAVVQKAAHGHVKAESKSHVSTLTDLOPYMRQFVAHL 780  
QY 743 PDRLTEVIASIIKPPONTYCVRYAVVQKAAHGHVKAESKSHVSTLTDLOPYMRQFVAHL 802

Db 781 QETSPLRDAVVTEQSSSLNEASSGLFDVFLRFWCHHAVIRKSKSVQCOGIPQGSILSTL 840  
QY 803 QETSPLRDAVVTEQSSSLNEASSGLFDVFLRFWCHHAVIRKSKSVQCOGIPQGSILSTL 862

Db 841 LCSLCYGDMMENKLFAGIRRDGLLLRLVDDFLVTLPHLTHAKTFLRFLVRGVPEYGCVVNL 900  
QY 863 LCSLCYGDMMENKLFAGIRRDGLLLRLVDDFLVTLPHLTHAKTFLRFLVRGVPEYGCVVNL 922

Db 901 RKTVVNFVPEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQDSYSSYARTSIRASLTF 960  
QY 923 RKTVVNFVPEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVSQDSYSSYARTSIRASLTF 982

Db 961 NRGFRAGRNMRKLFGLVLRKCHSLFLDLQVNSLOTVCNIIKILLQAYRFHACVLQLP 1020  
QY 983 NRGFRAGRNMRKLFGLVLRKCHSLFLDLQVNSLOTVCNIIKILLQAYRFHACVLQLP 1042

Db 1021 FHQQVWKNPTFFELRVISDTASLCYSTLKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1080  
QY 1043 FHQQVWKNPTFFELRVISDTASLCYSTLKAKNAGMSLGAKGAGPLPSEAVQWLCHQAFLL 1102

Db 1081 KLTRHRVTYVPLLSGLRTAQTSLSRKLPGTTLTALFAAANPALPSDFKTILD 1132  
QY 1103 KLTRHRVTYVPLLSGLRTAQTSLSRKLPGTTLTALFAAANPALPSDFKTILD 1154

RESULT 21  
ID US-08-974-549-611 STANDARD: PRT: 1154 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 611, Application US/08974549  
XX  
CC Sequence 611, Application US/08974549  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
CC NUMBER OF SEQUENCES: 726  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/974,549  
CC FILING DATE: 19-NOV-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph Ted  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002610US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 611:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1154 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: Protein  
CC LOCATION: 1..1154  
CC OTHER INFORMATION: /note= "fusion protein composed of hTERT  
CC OTHER INFORMATION: protein sequence, vector sequences, the  
CC OTHER INFORMATION: Myc epitope and His6 tag"  
CC SEQUENCE 1154 AA: 129603 MW: 6816882 CN:  
  
Query Match 98.28; Score 8465; DB 18; Length 1154;  
Best Local Similarity 100.08; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRLGPGQWRLVQSGDPAAFRALVAQCLVCPW 60  
QY |||||  
Db 23 MPAPRCRAVRSLLRSHYREVLPATFVRLGPGQWRLVQSGDPAAFRALVAQCLVCPW 82  
QY |||||  
Db 61 DARPPPAAPSFQVSCIKELVARVQLRCERGAKNVLAFCGALLDGAAGGPPPEAFTTSVR 120  
QY |||||  
Db 83 DARPPPAAPSFQVSCIKELVARVQLRCERGAKNVLAFCGALLDGAAGGPPPEAFTTSVR 142  
QY |||||  
Db 121 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYOLGA 180  
QY |||||  
Db 143 SYLPNTVTDALRGSGAWGLLRVGGDVLVLLARCALFVLVAPSCAYQVCGPPLYOLGA 202  
QY |||||  
Db 181 ATQARPPPHASGPRRLGGERANNSHVSREAGVPLGLPAPGARRRGSGASRLPLPKRPRR 240  
QY |||||  
Db 203 ATQARPPPHASGPRRLGGERANNSHVSREAGVPLGLPAPGARRRGSGASRLPLPKRPRR 262  
QY |||||  
Db 241 GAAPERTPVGGSWAHGPORTRGSDRGFCVVSPPARPAEATSLGALSGTRHSPSVG 300  
QY |||||  
Db 263 GAAPERTPVGGSWAHGPORTRGSDRGFCVVSPPARPAEATSLGALSGTRHSPSVG 322  
QY |||||  
Db 301 RQHHAGPPSTSRPRPMDTCPVYATKHFLLYSSGDKQLRFSLLSLRPLSGARLL 360  
QY |||||  
Db 323 RQHHAGPPSTSRPRPMDTCPVYATKHFLLYSSGDKQLRFSLLSLRPLSGARLL 382  
QY |||||  
Db 361 VETIFLGSRPMPGTPRRPLRPLQRYWQMPLELLIGNHAQCPYGVLLKTHCPRAAAT 420  
QY |||||  
Db 383 VETIFLGSRPMPGTPRRPLRPLQRYWQMPLELLIGNHAQCPYGVLLKTHCPRAAAT 442  
QY |||||

Db 421 PAAGVCAREKPOGSGVAAPPEEDTDPRLVQLLRQHSPPQVYGFVRACLRRLVPPGLWGS 480  
QY |||||  
Db 443 PAAGVCAREKPOGSGVAAPPEEDTDPRLVQLLRQHSPPQVYGFVRACLRRLVPPGLWGS 502  
QY |||||  
Db 481 RNREFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHRLREEL 540  
QY |||||  
Db 503 RNREFLRNTKFFISLGHAKLSLQELTWKMSVRDCAWLRSPGVGCVPAAEHRLREEL 562  
QY |||||  
Db 541 LAKFHLWMSVTVVVELLSFFVYVTTTFFOKNRLFFYKSWKSLQSIGIRQHLKRVQLE 600  
QY |||||  
Db 563 LAKFHLWMSVTVVVELLSFFVYVTTTFFOKNRLFFYKSWKSLQSIGIRQHLKRVQLE 622  
QY |||||  
Db 601 LSEAEVRQHREARPAALLTSRLRFIPKPDGLRPVNNMDYVVGARTFRREKRAELTSRKA 660  
QY |||||  
Db 623 LSEAEVRQHREARPAALLTSRLRFIPKPDGLRPVNNMDYVVGARTFRREKRAELTSRKA 682  
QY |||||  
Db 661 LFSVLNTERARRPGLLGASVLGDDIHRARWTFVLRVRAQDPPPELYFKYVDVTGAYDTI 720  
QY |||||  
Db 683 LFSVLNTERARRPGLLGASVLGDDIHRARWTFVLRVRAQDPPPELYFKYVDVTGAYDTI 742  
QY |||||  
Db 721 PQDLTEVIASIIKQNTYCVRRYAVVQKAHGHVKAFAKSHVSTLTDLQPYMRQFVAHL 780  
QY |||||  
Db 743 PQDLTEVIASIIKQNTYCVRRYAVVQKAHGHVKAFAKSHVSTLTDLQPYMRQFVAHL 802  
QY |||||  
Db 781 QETSPLRDVAVVIEQSSSLNEASSGLFDVFLRFWCHHVAIRKSYVQCQIGIPOGSLSTL 840  
QY |||||  
Db 803 QETSPLRDVAVVIEQSSSLNEASSGLFDVFLRFWCHHVAIRKSYVQCQIGIPOGSLSTL 862  
QY |||||  
Db 841 LCSLCYGDMEKLFAGIRRDGLLLRLVDVDFLLVTPHLLTHAKTFLRLTVRGVPEYGCWNL 900  
QY |||||  
Db 863 LCSLCYGDMEKLFAGIRRDGLLLRLVDVDFLLVTPHLLTHAKTFLRLTVRGVPEYGCWNL 922  
QY |||||  
Db 901 RKTVYNFPVEDEALGATFQVMPAHGLFPWCGLLDTRTLEVOSSDYSSYARTSIRASLTF 960  
QY |||||  
Db 923 RKTVYNFPVEDEALGATFQVMPAHGLFPWCGLLDTRTLEVOSSDYSSYARTSIRASLTF 982  
QY |||||  
Db 961 NRGFKAGRNMRKLFGLVKLCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1020  
QY |||||  
Db 983 NRGFKAGRNMRKLFGLVKLCHSLFLDLQVNSLQTVCTNIYKILLQAYRFHACVQLP 1042  
QY |||||  
Db 1021 FHOQWKNTEFRLVSDTASLCYSTLKAKNAGMSLGAKAGPLPSEAVQWLCHQAFLL 1080  
QY |||||  
Db 1043 FHOQWKNTEFRLVSDTASLCYSTLKAKNAGMSLGAKAGPLPSEAVQWLCHQAFLL 1102  
QY |||||  
Db 1081 KLTRHRTVYVPLLSLRTAQTLRSKLPGLTTLTAAANPALPSDFKTILD 1132  
QY |||||  
Db 1103 KLTRHRTVYVPLLSLRTAQTLRSKLPGLTTLTAAANPALPSDFKTILD 1154  
QY |||||

RESULT 22  
ID US-08-911-312-45 STANDARD; PRT: 1154 AA.  
XX  
AC xxxxxx  
XX  
XX  
XX  
DE Sequence 45, Application US/08911312  
XX  
XX Sequence 45, Application US/08911312  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco

CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911,312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 23-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-002500US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 45:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1154 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS:  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1154 AA; 129603 MW; 6816882 CN;  
  
Query Match 98.2%; Score 8465; DB 13; Length 1154;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 MPAPRCRAVRSLLRSHRYEVLPLATFVRRLLGPGQWRLVQGRDPAAPRALYVAQCIVCPW 60  
Qy 23 MPAPRCRAVRSLLRSHRYEVLPLATFVRRLLGPGQWRLVQGRDPAAPRALYVAQCIVCPW 82  
  
Db 61 DARPPPAAPSPFQVSCLELVARLQRCERGAKNVLAFFGALLDGGAGGPPPEAFTTSVR 120  
Qy 83 DARPPPAAPSPFQVSCLELVARLQRCERGAKNVLAFFGALLDGGAGGPPPEAFTTSVR 142  
  
Db 121 SYLPTNTVDALRGSGAGWGLLLRRYGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180  
Qy 143 SYLPTNTVDALRGSGAGWGLLLRRYGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202  
  
Db 181 ATQARPPHSGPRRRLGRCERAWNHSREAGVPLGLPAGARRRGGSASRSLPLPKRPRR 240  
Qy 203 ATQARPPHSGPRRRLGRCERAWNHSREAGVPLGLPAGARRRGGSASRSLPLPKRPRR 262  
  
Db 241 GAAPEPRTPVGCGSWAHPGTRGSPDRGFCVSPARPABEATSLGALSGTRHSHPSVG 300  
Qy 263 GAAPEPRTPVGCGSWAHPGTRGSPDRGFCVSPARPABEATSLGALSGTRHSHPSVG 322

301 ROHHAGPSTSRPPRWDTPCPCPPVYAEKHFLLYSKGDEQLRPSFLSSLRPSLTGARRL 360  
323 ROHHAGPSTSRPPRWDTPCPCPPVYAEKHFLLYSKGDEQLRPSFLSSLRPSLTGARRL 382  
361 VEIIFGSRPWPCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAVT 420  
383 VEIIFGSRPWPCTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPPLRAVT 442  
421 PAAGVCAREKPOGSVAAPAEEDTDPRRLVQLLRQHSPPQVYGFVRACLRLRPPGLWGS 480  
443 PAAGVCAREKPOGSVAAPAEEDTDPRRLVQLLRQHSPPQVYGFVRACLRLRPPGLWGS 502  
481 RHNERFLRNTKKEFISLGKHAQLSLQELTWKMSVRDCAWLRSPGCVCPAAEHLREBI 540  
503 RHNERFLRNTKKEFISLGKHAQLSLQELTWKMSVRDCAWLRSPGCVCPAAEHLREBI 562  
541 LAKFLHLMMSVYVVELLSFFYVTFQKNRLFYFKSVWSKLSIGIRQHLKRVOLRE 600  
563 LAKFLHLMMSVYVVELLSFFYVTFQKNRLFYFKSVWSKLSIGIRQHLKRVOLRE 622  
601 LSEAEVRQHREARPALTSRLRFIPKPDGLRPVNMMDYVVGARTFRREKRAERLTSRVKA 660  
623 LSEAEVRQHREARPALTSRLRFIPKPDGLRPVNMMDYVVGARTFRREKRAERLTSRVKA 682  
661 LFSVLNTERARRPGLLGASVGLGDDIHRARWTFVLVRADQDPPPELVFVVDVTGAYDTI 720  
683 LFSVLNTERARRPGLLGASVGLGDDIHRARWTFVLVRADQDPPPELVFVVDVTGAYDTI 742  
721 PQDLTEVIASIIKPNQTYCVRVAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 780  
743 PQDLTEVIASIIKPNQTYCVRVAVVQKAAHGHVRKAFKSHVSTLTDLPYMRQFVAHL 802  
781 QETSPLRDAVVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVOCQIGPOGSIILSTL 840  
803 QETSPLRDAVVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVOCQIGPOGSIILSTL 862  
841 LCSLCYGDMEKLFAGIRRDGGLLLRLVDDFLLVTPHLLTHAKTELRTLVRGVPYGVVNL 900  
863 LCSLCYGDMEKLFAGIRRDGGLLLRLVDDFLLVTPHLLTHAKTELRTLVRGVPYGVVNL 922  
901 RKTWNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVQSDSYASRTSIRASLTF 960  
923 RKTWNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVQSDSYASRTSIRASLTF 982  
961 NRGFKAGRNMRKLFGLVRLKCHSLFDLQVNSLQTVCTNIIYKILLQAYRFHACVQLP 1020  
983 NRGFKAGRNMRKLFGLVRLKCHSLFDLQVNSLQTVCTNIIYKILLQAYRFHACVQLP 1042  
1021 FHOQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAAGPLPSEAYOWLCHQAFLL 1080  
1043 FHOQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAAGPLPSEAYOWLCHQAFLL 1102  
1081 KLFRHRYTYVPLLSLRTAQTLRSKLPGLTTLTALEAANPALPSDFKTILD 1132  
1103 KLFRHRYTYVPLLSLRTAQTLRSKLPGLTTLTALEAANPALPSDFKTILD 1154

RESULT 23  
ID US-09-026-981-3 STANDARD: PRT; 1132 AA.  
XX AC xxxxxx  
XX XX  
DT XX  
XX XX  
DE Sequence 3, Application US/09026981  
XX  
CC Sequence 3, Application US/09026981  
CC GENERAL INFORMATION:  
CC APPLICANT: Counter, Christopher M.  
CC APPLICANT: Meyerson, Matthew  
CC APPLICANT: Weinberg, Robert A.  
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and  
CC NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,981  
FILING DATE: 20-FEB-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/064,322  
FILING DATE: 30-OCT-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/055,762  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/054,549  
FILING DATE: 01-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,151  
FILING DATE: 20-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/038,750  
FILING DATE: 20-FEB-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WH197-11p4AM  
TELEPHONE: 781-861-6240  
TELEFAX: 781-861-9540

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 1132 AA; 126937 MW; 6590273 CN;  
Query Match 98.1%; Score 8459; DB 14; Length 1132;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MPAPRCRAVRLSLRSHYREVLPATFVRRILGPOGWRILVQRGDPAAFRALVAQCLVCPW 60

23 MPAPRCRAVRLSLRSHYREVLPATFVRRILGPOGWRILVQRGDPAAFRALVAQCLVCPW 82

61 DARPPAPSPRQVSKLVELVARVQLRCERGAKNVLAFGFALLDARGGPEAFTTSVR 120

83 DARPPAPSPRQVSKLVELVARVQLRCERGAKNVLAFGFALLDARGGPEAFTTSVR 142

121 SYLPTNTDALRGSGANGLLRRVDDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180

143 SYLPTNTDALRGSGANGLLRRVDDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 202

181 ATQARPPPHASGPRRLRCERAWNHSREAGVPLGLPAPGARRRGGASRSRLPLPKRPRR 240

203 ATQARPPPHASGPRRLRCERAWNHSREAGVPLGLPAPGARRRGGASRSRLPLPKRPRR 262

241 GAAPERTPVGQSWAHPGTRGSDRGFCVSPARPAEATSLGALSSTRHSHPSVG 300

263 GAAPERTPVGQSWAHPGTRGSDRGFCVSPARPAEATSLGALSSTRHSHPSVG 322

301 ROHAGPPSTSRPPWDTPCPVYAEKHFYSSGDKQELRPSFLSSLRPSLTGARRL 360

|||||

QY 323 ROHAGPPSTSRPPWDTPCPVYAEKHFYSSGDKQELRPSFLSSLRPSLTGARRL 382  
Db 361 VETIFLGSREPMFGTTPRRLPRLPQRYWQMRPLFLELIGNHAQCPYGYLLKTHCPRAAVT 420  
CC |||||||  
CC |||||||  
QY 383 VETIFLGSREPMFGTTPRRLPRLPQRYWQMRPLFLELIGNHAQCPYGYLLKTHCPRAAVT 442  
Db 421 PAAGVCAREKPOGSVAAPPEEDTPRRLVOLLROHSSPMQVYGFVRACLRRLVPPGLWGS 480  
CC |||||||  
CC |||||||  
QY 443 PAAGVCAREKPOGSVAAPPEEDTPRRLVOLLROHSSPMQVYGFVRACLRRLVPPGLWGS 502  
Db 481 RHNERFLRNTKPFISLGHAKLSLOELTWKMSVRCGLRRSPGVCVPAAEHRLREEI 540  
CC |||||||  
CC |||||||  
QY 503 RHNERFLRNTKPFISLGHAKLSLOELTWKMSVRCGLRRSPGVCVPAAEHRLREEI 562  
Db 541 LAKFLHLMMSVYVVELLSRFFYVTTETFOKNRLFYFKSVWSKLQSIGIRHKLKRVOLRE 600  
CC |||||||  
CC |||||||  
QY 563 LAKFLHLMMSVYVVELLSRFFYVTTETFOKNRLFYFKSVWSKLQSIGIRHKLKRVOLRE 622  
Db 601 LSAEVRQHREARPALITSRLRIFPKPDGLRPIVNDYVVGARTFRREKRAERLTSRVKA 660  
CC |||||||  
CC |||||||  
QY 623 LSAEVRQHREARPALITSRLRIFPKPDGLRPIVNDYVVGARTFRREKRAERLTSRVKA 682  
Db 661 LFSVLNTERARRPGLLGASVIGLDDIHRAWRTFVLVRAODPPPELYFVKVDVTGAYDTI 720  
CC |||||||  
CC |||||||  
QY 683 LFSVLNTERARRPGLLGASVIGLDDIHRAWRTFVLVRAODPPPELYFVKVDVTGAYDTI 742  
Db 721 PODRLTEVIASIIKPQNTYCVRRYAVVQAAHGHVRAKFSHSTLTLDLQPMRQFVAHL 780  
CC |||||||  
CC |||||||  
QY 743 PODRLTEVIASIIKPQNTYCVRRYAVVQAAHGHVRAKFSHSTLTLDLQPMRQFVAHL 802  
Db 781 QETSPLRDADVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIGPOGSIILSTL 840  
CC |||||||  
CC |||||||  
QY 803 QETSPLRDADVIEQSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVQCQIGPOGSIILSTL 862  
Db 841 LCSLCYGDMEKLFAGLRDGLLLRLVDDLLVTPHLLTHAKTFLRTLVRGVPEYGCVVNL 900  
CC |||||||  
CC |||||||  
QY 863 LCSLCYGDMEKLFAGLRDGLLLRLVDDLLVTPHLLTHAKTFLRTLVRGVPEYGCVVNL 922  
Db 901 RKTVMNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVDSDYSSYARTSIRASLTF 960  
CC |||||||  
CC |||||||  
QY 923 RKTVMNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRTLEVDSDYSSYARTSIRASLTF 982  
Db 961 NRGFKAGRNMRKLFGLVRLKCHSLFDLDQVNSLQTVCTNIYKILLQAYRFHACVLQLP 1020  
CC |||||||  
CC |||||||  
QY 983 NRGFKAGRNMRKLFGLVRLKCHSLFDLDQVNSLQTVCTNIYKILLQAYRFHACVLQLP 1042  
Db 1021 FHOQVKNPTEFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHOAFL 1080  
CC |||||||  
CC |||||||  
QY 1043 FHOQVKNPTEFFLRVSDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVOWLCHOAFL 1102  
Db 1081 KLTRHRVTVYVPLGSLRTAQTLRSRLPGTTLTALAAANPALPSDFKTIID 1132  
CC |||||||  
CC |||||||  
QY 1103 KLTRHRVTVYVPLGSLRTAQTLRSRLPGTTLTALAAANPALPSDFKTIID 1154

RESULT 24

ID US-08-974-549-344 STANDARD; PRT; 1132 AA.

XX

AC

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 344, Application US/08974549

Sequence 344, Application US/08974549

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.





\*\*\*\*\*

MPERCH\_PP

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 24 08:00:10 1998; MasPar time 12.30 Seconds  
Tabular output not generated. 664.418 Million cell updates/sec

Title: >US-08-951-733-20  
Description: (1-1154) from US08951733.pep  
Perfect Score: 8624  
Sequence: 1 HASGQRCVLLRTWELAPAT.....TALEAANPALPSDFKTLTD 1154

Scoring table: PAM 150  
Gap 11

Searched: 77309 seqs, 7078906 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 37.265; Variance 182.038; scale 0.205

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	118	1.4	399	1	US-08-147- Sequence 27, Applicati	2.06e+00
2	118	1.4	399	1	US-08-447- Sequence 27, Applicati	2.06e+00
3	118	1.4	399	2	PCT-US93-0 Sequence 23, Applicati	2.06e+00
4	118	1.4	399	1	US-08-278- Sequence 23, Applicati	2.06e+00
5	118	1.4	399	1	US-08-643- Sequence 23, Applicati	2.06e+00
6	118	1.4	399	1	US-07-901- Sequence 13, Applicati	2.06e+00
7	118	1.4	399	1	US-08-643- Sequence 23, Applicati	2.06e+00
8	118	1.4	399	1	US-08-479- Sequence 10, Applicati	2.06e+00
9	118	1.4	399	1	US-08-155- Sequence 23, Applicati	2.06e+00
10	118	1.4	399	2	PCT-US93-0 Sequence 23, Applicati	2.06e+00
11	118	1.4	399	2	PCT-US93-0 Sequence 13, Applicati	2.06e+00
12	118	1.4	399	2	PCT-US93-1 Sequence 10, Applicati	2.06e+00
13	118	1.4	399	2	PCT-US93-0 Sequence 23, Applicati	2.06e+00
14	118	1.4	399	1	US-08-462- Sequence 23, Applicati	2.06e+00
15	118	1.4	399	2	PCT-US93-0 Sequence 8, Applicatio	2.06e+00
16	118	1.4	399	2	US-08-480- Sequence 23, Applicati	2.06e+00
17	118	1.4	399	1	PCT-US93-0 Sequence 10, Applicati	2.06e+00
18	118	1.4	399	2	PCT-US93-0 Sequence 23, Applicati	2.06e+00
19	118	1.4	399	1	US-08-406- Sequence 23, Applicati	2.06e+00
20	118	1.4	399	1	US-08-451- Sequence 23, Applicati	2.06e+00
21	118	1.4	399	2	PCT-US93-0 Sequence 23, Applicati	2.06e+00
22	118	1.4	399	2	PCT-US91-0 Sequence 6, Applicatio	2.06e+00
23	125	1.4	402	1	US-07-841- Sequence 29, Applicati	7.05e-01

24	125	1.4	402	1	US-08-643- Sequence 21, Applicati	7.05e-01
25	125	1.4	402	1	US-08-451- Sequence 21, Applicati	7.05e-01
26	125	1.4	402	1	US-08-462- Sequence 21, Applicati	7.05e-01
27	125	1.4	402	2	PCT-US93-0 Sequence 21, Applicati	7.05e-01
28	125	1.4	402	1	US-08-206- Sequence 4, Applicati	7.05e-01
29	125	1.4	402	1	US-08-155- Sequence 21, Applicati	7.05e-01
30	125	1.4	402	2	PCT-US93-0 Sequence 21, Applicati	7.05e-01
31	125	1.4	402	1	US-08-480- Sequence 8, Applicatio	7.05e-01
32	125	1.4	402	1	US-07-901- Sequence 11, Applicati	7.05e-01
33	125	1.4	402	1	US-08-147- Sequence 29, Applicati	7.05e-01
34	125	1.4	402	1	US-08-643- Sequence 21, Applicati	7.05e-01
35	125	1.4	402	1	US-08-406- Sequence 21, Applicati	7.05e-01
36	125	1.4	402	1	US-08-479- Sequence 8, Applicatio	7.05e-01
37	125	1.4	402	1	US-08-278- Sequence 21, Applicati	7.05e-01
38	125	1.4	402	1	US-08-447- Sequence 29, Applicati	7.05e-01
39	125	1.4	402	2	PCT-US93-0 Sequence 21, Applicati	7.05e-01
40	125	1.4	402	2	PCT-US93-0 Sequence 21, Applicati	7.05e-01
41	125	1.4	402	2	PCT-US93-0 Sequence 21, Applicati	7.05e-01
42	125	1.4	402	2	PCT-US93-1 Sequence 8, Applicatio	7.05e-01
43	125	1.4	402	2	PCT-US92-0 Sequence 21, Applicati	7.05e-01
44	125	1.4	402	2	PCT-US93-0 Sequence 11, Applicati	7.05e-01
45	125	1.4	402	2	PCT-US93-0 Sequence 6, Applicatio	7.05e-01

ALIGNMENTS

RESULT 1  
ID US-08-147-023-27 STANDARD; PRT; 399 AA.  
XX  
AC xxxxxx  
XX  
DT

Sequence 27, Application US/08147023

Sequence 27, Application US/08147023

Patent No. 5468845

GENERAL INFORMATION:

APPLICANT: OPPERMAN, HERMANN

APPLICANT: OZKAYNAK, ENGIN

APPLICANT: KUBERASAMPATH, THANGAVEL

APPLICANT: RUEGER, DAVID C.

APPLICANT: PANG, ROY H. L.

TITLE OF INVENTION: OSTEOGENIC DEVICES

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/147,023

FILING DATE: 21-FEB-1992

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 810,560

FILING DATE: 20-DEC-1991

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 827,052

FILING DATE: 28-JAN-1992

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 660,162

FILING DATE: 22-FEB-1991

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 621,988

FILING DATE: 04-DEC-1990





CC TELEPHONE: 617/248-7000  
CC TELEFAX: 617/248-7100  
CC INFORMATION FOR SEQ ID NO: 27:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 1; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
  
Db 12 GLALCALGGHGPRPP-HTC-PQRRLG-ARERRDMQREILAVLGLPGR-PRRPAQPAAR 67  
QY 194 GPPLYQLGAATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRS 253  
Db 68 QP 69  
QY 254 LP 255  
  
RESULT 3  
ID PCT-US92-01968-23 STANDARD; PRT; 399 AA.  
XX AC xxxxxx  
XX DT  
XX DE  
XX DE Sequence 23, Application PC/TUS9201968  
XX DE Sequence 23, Application PC/TUS9201968  
CC GENERAL INFORMATION:  
CC APPLICANT: COHEN, CHARLES M  
CC APPLICANT: KUBERASAMPATH, THANGAVEL  
CC APPLICANT: PANG, ROY HL  
CC APPLICANT: OPPERMANN, HERMAN  
CC APPLICANT: RUEGER, DAVID C  
CC TITLE OF INVENTION: PROTEIN-INDUCED MORPHOGENESIS  
CC NUMBER OF SEQUENCES: 23  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: TESTA HURWITZ & THIBEAULT  
CC STREET: EXCHANGE PLACE 53 STATE STREET  
CC CITY: BOSTON  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02109  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US92/01968  
CC FILING DATE: 19920311  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 667,274  
CC FILING DATE: 11-MAR-1991  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 752,764  
CC FILING DATE: 30-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: PITCHER ESO, EDMUND R  
CC REGISTRATION NUMBER: 27,829  
CC REFERENCE/DOCKET NUMBER: CRP-052PC  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617/248-7000  
CC INFORMATION FOR SEQ ID NO: 23:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: AMINO ACID

CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 2; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
  
Db 12 GLALCALGGHGPRPP-HTC-PQRRLG-ARERRDMQREILAVLGLPGR-PRRPAQPAAR 67  
QY 194 GPPLYQLGAATQARPPPHASGPRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRS 253  
Db 68 QP 69  
QY 254 LP 255  
  
RESULT 4  
ID US-08-278-729A-23 STANDARD; PRT; 399 AA.  
XX AC xxxxxx  
XX DT  
XX DE  
XX DE Sequence 23, Application US/08278729A  
XX DE Sequence 23, Application US/08278729A  
CC Patent No. 5650276  
CC GENERAL INFORMATION:  
CC APPLICANT: SMART, JOHN  
CC APPLICANT: OPPERMANN, HERMAN  
CC APPLICANT: OZKAYNAK, ENGIN  
CC APPLICANT: KUBERASAMPATH, THANGAVEL  
CC APPLICANT: RUEGER, DAVID C  
CC APPLICANT: PANG, ROY H.L.  
CC APPLICANT: COHEN, CHARLES M.  
CC TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
CC NUMBER OF SEQUENCES: 33  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
CC STREET: 45 SOUTH STREET  
CC CITY: HOPKINTON  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 01748  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/278,729A  
CC FILING DATE: 20-JUL-1994  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: PITCHER Esq., EDMUND R.  
CC REGISTRATION NUMBER: 27,829  
CC REFERENCE/DOCKET NUMBER: CRP-058CPFW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (508) 435-9001  
CC TELEFAX: (508) 435-6951  
CC INFORMATION FOR SEQ ID NO: 23:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 1; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;





CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/155,343A  
CC FILING DATE: 15-NOV-1993  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: FENTON ESQ., GILLIAN M.  
CC REGISTRATION NUMBER: 36,508  
CC REFERENCE/DOCKET NUMBER: CRP-067FW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617) 248-7560  
CC TELEFAX: (617) 248-7100  
CC INFORMATION FOR SEQ ID NO: 23:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 1; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
  
Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMQREILAVLGLPGR-PPRAQPAAR 67  
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253  
  
Db 68 QP 69  
QY 254 LP 255  
  
RESULT 10  
ID PCT-US93-07231-23 STANDARD; PRT; 399 AA.  
XX xxxxxx  
DT  
XX  
DE Sequence 23, Application PC/TUS9307231  
XX  
XX Sequence 23, Application PC/TUS9307231  
CC GENERAL INFORMATION:  
CC APPLICANT:  
CC TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND  
CC NUMBER OF SEQUENCES: 33  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.  
CC STREET: 35 SOUTH STREET  
CC CITY: HOPKINTON  
CC STATE: MASSACHUSETTS  
CC COUNTRY: USA  
CC ZIP: 01748  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/07231  
CC FILING DATE: 19930729  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: KELLEY, ROBIN D.  
CC REGISTRATION NUMBER: 34,637  
CC REFERENCE/DOCKET NUMBER: CRP-070  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617/248-7000

CC TELEFAX: 617/248-7100  
CC INFORMATION FOR SEQ ID NO: 23:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 2; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
  
Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMQREILAVLGLPGR-PPRAQPAAR 67  
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253  
  
Db 68 QP 69  
QY 254 LP 255  
  
RESULT 11  
ID PCT-US93-05446-13 STANDARD; PRT; 399 AA.  
XX xxxxxx  
DT  
XX  
DE Sequence 13, Application PC/TUS9305446  
XX  
XX Sequence 13, Application PC/TUS9305446  
CC GENERAL INFORMATION:  
CC APPLICANT:  
CC TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED  
CC NUMBER OF SEQUENCES: 22  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.  
CC STREET: 35 SOUTH STREET  
CC CITY: HOPKINTON  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 01748  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/05446  
CC FILING DATE: 19930608  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: PITCHER ESQ, EDMUND R  
CC REGISTRATION NUMBER: 27,829  
CC REFERENCE/DOCKET NUMBER: STR-057  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617/248-7000  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 399 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;  
  
Query Match 1.4%; Score 118; DB 2; Length 399;  
Best Local Similarity 38.7%; Pred. No. 2.06e+00;  
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;  
  
Db 12 GLALCALGGGHGPRPP-HTC-PQRLG-ARERDMQREILAVLGLPGR-PPRAQPAAR 67  
QY 194 GPLYQLGAATQARPPPHASGPRRLGCRANWHSVREAGVPLGLPAPGARRGGSASRS 253





\*\*\*\*\*  
WILEY  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Blocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 08:54:27 1998; MasPar time 9229.89 Seconds  
Tabular output not generated.  
680.010 Million cell updates/sec

Title: >US-08-951-733-13  
(1-2848) from US08951733.seq (1 of 3)  
Perfect Score: 2848  
N.A. Sequence: 1 CACGGCTCCGGCGACGCTG.....GATGCCGCCACGCCCVAT 2848  
Comp: GTGCGCAGCGCGTCCGCAC.....CTACGCCGGGTGCGCGATA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hun1 5:em\_hun2 6:em\_in  
7:em\_or 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_v1

Database: genbank107  
15:gb\_ba1 16:gb\_ba2 17:gb\_htg 18:gb\_in 19:gb\_or 20:gb\_ov  
21:gb\_pat 22:gb\_ph 23:gb\_pl1 24:gb\_pl2 25:gb\_pr1  
26:gb\_pr2 27:gb\_pr3 28:gb\_ro 29:gb\_st 30:gb\_sts 31:gb\_sy  
32:gb\_un 33:gb\_v1

Statistics: Mean 11.679; Variance 6.056; scale 1.928

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	2837	99.6	4015 27 AF015950 Homo sapiens telomeras 0.00e+00
2	2824	99.2	8960 31 AF043739 Synthetic construct hu 0.00e+00
3	2823	99.1	4027 27 AF018167 Homo sapiens telomeras 0.00e+00
4	825	29.0	3426 28 AF051911 Mus musculus telomeras 0.00e+00
5	823	28.9	3369 28 AF073311 Mus musculus telomeras 0.00e+00
6	129	4.5	201 28 AF029235 Mus musculus telomeras 1.63e-71
7	67	2.4	7218 21 I66494 Sequence 14 from paten 1.13e-25
8	49	1.7	7218 21 I66494 Sequence 14 from paten 1.72e-13
9	37	1.3	74371 26 AC005369 Homo sapiens chromosom 5.03e-06
10	34	1.2	215 21 I28278 Sequence 5 from patent 2.75e-04
11	35	1.2	215 21 I28278 Sequence 5 from patent 7.37e-05
12	34	1.2	74371 26 AC005369 Homo sapiens chromosom 2.75e-04

13	30	1.1	201 21	A10158	S.griseus gene for pre	4.44e-02
14	30	1.1	201 21	A10161	Synthetic DNA for prep	4.44e-02
15	30	1.1	201 21	A10162	Synthetic DNA for prep	4.44e-02
16	30	1.1	201 21	A10159	S.griseus gene for pre	4.44e-02
17	30	1.1	201 21	E04076	gDNA encoding envelope	4.44e-02
18	31	1.1	216021	26 HUAC004787	Homo sapiens Chromosom	1.28e-02
19	28	1.0	16623	19 HU92534	Mustela vison microsat	4.95e-01
20	29	1.0	6011	33 HS4RSIRZ2T	Epstein-Barr virus (wi	1.50e-01
21	29	1.0	15807	15 MEU72662	Methylobacterium extor	4.95e-01
22	28	1.0	31812	24 SPUNK4	S.pombe chromosome I c	4.95e-01
23	28	1.0	39073	23 SPAC2E11	S.pombe chromosome I c	4.95e-01
24	29	1.0	76094	25 HSL79D3A	Human DNA sequence fro	1.50e-01
25	27	0.9	565	21 E04076	gDNA encoding envelope	1.59e+00
26	27	0.9	904	33 HS4BHLF	Epstein-Barr virus Bam	1.59e+00
27	26	0.9	1341	20 CHKD1LRB	Gallus domesticus DI-1	4.98e+00
28	26	0.9	2926	15 SNAPROBA	S.marcescens probA ope	4.98e+00
29	26	0.9	2926	15 SMPROBAG	S.marcescens probA gen	4.98e+00
30	26	0.9	4767	28 AF034863	Rattus norvegicus syna	4.98e+00
31	27	0.9	5050	33 HS4U2IR2A	Epstein-Barr virus (B9	1.59e+00
32	27	0.9	5400	15 SCACTVA	S.coelicolor 6 actVA r	1.59e+00
33	26	0.9	12949	16 AF011922	Pseudomonas aeruginosa	4.98e+00
34	26	0.9	35980	16 MTY20H10	Mycobacterium tubercul	1.59e+00
35	27	0.9	43034	26 HSN5H5	Human DNA sequence fro	4.98e+00
36	27	0.9	64078	18 AC004365	Drosophila melanogaste	1.59e+00
37	26	0.9	85132	26 U82668	Homo sapiens shox gene	4.98e+00
38	26	0.9	123491	17 HS126A5	Human DNA sequence ***	4.98e+00
39	26	0.9	125815	17 HS436C18	Human DNA sequence ***	4.98e+00
40	26	0.9	138706	17 AC004946	*** SEQUENCING IN PROG	4.98e+00
41	26	0.9	147216	27 AF093117	Homo sapiens chromosom	4.98e+00
42	26	0.9	147542	17 HS754E20	Human DNA sequence ***	4.98e+00
43	26	0.9	172281	33 EBV	Epstein-Barr virus (EB	1.59e+00
44	27	0.9	184113	33 HS4B958RAJ	Epstein-Barr virus, ar	1.59e+00
45	27	0.9	184113	33 HS4B958RAJ	Epstein-Barr virus, ar	1.59e+00

ALIGNMENTS

RESULT	1	AF015950	4015 bp	mRNA	PRI	16-AUG-1997
LOCUS						
DEFINITION						
ACCESSION						
NID						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
Location/Qualifiers						
1. .4015						
/organism="Homo sapiens"						
/db_xref="taxon:9606"						
/tissue_type="kidney"						
/map="5p"						
/dev_stage="embryo"						
/chromosome="5"						
1. .4015						
/gene="hTERT"						
56. .3454						
/gene="hTERT"						
/codon_start=1						

/product="telomerase reverse transcriptase"  
/db\_xref="pid:g2330017"

AAAFALVAQPCVCFVFWPPAPPAFSPFQVSLKELVARVQLRCEKAGKNVLAFGFA  
LIDGARGPPFAFTTSVRSYLPNTVTDALRGSGAWGLLRVGDVLLHLLARCALFV  
LVAPSCAYCGPPLYLQGAATQARPAPHASPPRRLLGGERAWNHSVREAGVPLGLPA  
PGARRGSGASRLPLPRPRRGAPEPERTPVGGSWAHPGRTGSPDSFCVVPSPA  
RPAEATSEALGSTRSHSPVSGRHHAGPPSSRPPRPMDTFCPPVYAEKHLFLYS  
SGDKQLRPLSSLSRLTGARLVETIFLGRPMWPGTPRRLPRLPQRYWQMLPL  
FLELGNHAQCPYGYLLTHGCLRAAVTAAAGVAREKPOGSAAPBEDDPRRLVQ  
LHRHSSPWQYTFVFRACRLRLVPLGLWSRHNRRFLNKKFISLGHAKLSLOEL  
TWKMSVRDCAWLRRSPGVGCPAAERHLEILAKFLHMLSVVYVLEHRSFFVYET  
TFQKNLFYKSWKSLQSIGIRQHLRVQLRELSEAEVQHRERAPALLTSRLFI  
KPDGLRPIVNDYVYGARTFEREARELTSRVKALFSVNLNERRAPRPLGLASVLG  
LONIRVETFLVRVADQPPPELYFVVDYTGAYDTIPQDLRETEVIAISLIPQNTYC  
VRRYAVQNAKCHVRKASHVSLTDLPYMRQFVAHLQETSPDRADVIEGSSSL  
NEASGLDFVLRFMCHAVIRKGSYVQCOIPQGSILSLILSLCYDMENLEFAG  
IRRDGLLILRLDVLVPLHTHAKTFLRTLVRGVEYGVVNLKRTVNPVPRDEAL  
GDTAFVOMPAHGLPWCGLLDDTRTLEVQSDYSVARTSIRASLTFNFGKAGWRMR  
KLFGLRLKCHSLFLDLQVNSIQVTCNIYKILLLOAYRFHACVLQPLFFHQVWKNPT  
FPLRVISDASICSILKAKNAGMSLGAAGAPLPSSEAVOMLCHQAFLLKLTRHRT  
YVPLGLSRLTAQTLQSLRPLGTTTALFAANPALPSDFKILD"

BASE COUNT 663 a 1363 c 1275 g 714 t

Query Match 99.6%; Score 2837; DB 27; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	661	CTGGAAACCATAGCGTCAGGAGGCGCGGGTCCCTCGGGCTTGCAGCCCGGGGTGCGAG	720
Qy	672	CTGGAAACCATAGCGTCAGGAGGCGCGGGTCCCTCGGGCTTGCAGCCCGGGGTGCGAG	731
Db	721	GAGGCGCGGGGCACTGCCAGCCGAAGTCTCCGCTTGCCCAAGAGGCCAGCGTGGCGC	780
Qy	732	GAGGCGCGGGGCACTGCCAGCCGAAGTCTCCGCTTGCCCAAGAGGCCAGCGTGGCGC	791
Db	781	TGCCCCCTGAGCGGAGCGAGCGCCGCTTGGGACAGGGGTCTCTGGGCCCCACCCGGGCGAGC	840
Qy	792	TGCCCCCTGAGCGGAGCGAGCGCCGCTTGGGACAGGGGTCTCTGGGCCCCACCCGGGCGAGC	851
Db	841	CGGTGACCGAGTACCCTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
Qy	852	CGGTGACCGAGTACCCTGT	911
Db	901	CACCTCTTTTGAGGGTGGCTCTCTGGACCGGCGACCTCCACCCATCCGTGGGCGCGCA	960
Qy	912	CACCTCTTTTGAGGGTGGCTCTCTGGACCGGCGACCTCCACCCATCCGTGGGCGCGCA	971
Db	961	GCACACGCGGCGCCCATCCACATCCGCGCCACACCTCCCTGGGACACGCTTGTTC	1020
Qy	972	GCACACGCGGCGCCCATCCACATCCGCGCCACACCTCCCTGGGACACGCTTGTTC	1031
Db	1021	CCGGGTGACCGGAGCAAGCACTTCTCTACTCTCAGGCGACAAAGAGGACAGCTGCG	1080
Qy	1032	CCGGGTGACCGGAGCAAGCACTTCTCTACTCTCAGGCGACAAAGAGGACAGCTGCG	1091
Db	1081	GCCCTCTCTCTACTCTCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1140
Qy	1092	GCCCTCTCTCTACTCTCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1151
Db	1141	GACCATCTTTCTGGGTTCCAGGCGCTTGATCCAGGAGGACTCCCGCGAGGTTGCCCGGCT	1200
Qy	1152	GACCATCTTTCTGGGTTCCAGGCGCTTGATCCAGGAGGACTCCCGCGAGGTTGCCCGGCT	1211
Db	1201	CGCCGAGCGCTACTGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1260
Qy	1212	CGCCGAGCGCTACTGGCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1271
Db	1261	GTGCGGCTACGCGGCTCTCTCAAGACGCACTGCGCGCTGGGAGCTGGGCTACCCGAGC	1320
Qy	1272	GTGCGGCTACGCGGCTCTCTCAAGACGCACTGCGCGCTGGGAGCTGGGCTACCCGAGC	1331
Db	1321	AGCGGCTGTCTGTCCGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1380
Qy	1332	AGCGGCTGTCTGTCCGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1391
Db	1381	CACAGACCGCGCTCGCTGT	1440
Qy	1392	CACAGACCGCGCTCGCTGT	1451
Db	1441	CGGCTCTGTGGGCGGCTCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1500
Qy	1452	CGGCTCTGTGGGCGGCTCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1511
Db	1501	CAACGAACCGCGCTCTCTCAGGAACACCAAGATTCATCTCCCTGGGAGCATGCGCA	1560
Qy	1512	CAACGAACCGCGCTCTCTCAGGAACACCAAGATTCATCTCCCTGGGAGCATGCGCA	1571
Db	1561	GCTCTCGCTGAGAGCTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAG	1620
Qy	1572	GCTCTCGCTGAGAGCTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAG	1631
Db	1621	GAGCCGAGGGTGGCTGTGTCTCGGCGCGGAGACCGCTCTGCTGAGGAGATCTTGGC	1680
Qy	1632	GAGCCGAGGGTGGCTGTGTCTCGGCGCGGAGACCGCTCTGCTGAGGAGATCTTGGC	1691
Db	1681	CAAGTTCTGCACTGGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1740
Qy	1692	CAAGTTCTGCACTGGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1751
Db	1741	TGTACCGGAGACCGCTTTTCAAAAGAACAGGCTCTTTTCTACCGGAGAGTGTCTGGAG	1800



```

1752 TGTACGGAGACCGTTCATAAAGACAGGCTCTTTTCTACCGAGAGCTGCTGGAG 1811
1801 CAAGTTGCAAGCATTTGAATCAGACAGCATTTGAAGAGGGTGCAGCTGCGGAGCTGTC 1860
1812 CAAGTTGCAAGCATTTGAATCAGACAGCATTTGAAGAGGGTGCAGCTGCGGAGCTGTC 1871
1861 GGAAGCAGAGGTCAGCAGCATTCGGAGAGCCAGGCGCCCTGCTGAGCTCCAGACTCCG 1920
1872 GGAAGCAGAGGTCAGCAGCATTCGGAGAGCCAGGCGCCCTGCTGAGCTCCAGACTCCG 1931
1921 CTTTCATCCCAAGCCTGACGGCTCGCGCGATTTGAACTGAGTACGTCTGCTGGAGC 1980
1932 CTTTCATCCCAAGCCTGACGGCTCGCGCGATTTGAACTGAGTACGTCTGCTGGAGC 1991
1981 CAGAAGCTTCCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGGCACTGTT 2040
1992 CAGAAGCTTCCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGGCACTGTT 2051
2041 CAGCGTGTCACTACGAGCGGCGCGCGCGCGCGCGCTCTGCGGCGCTCTGCTGGG 2100
2052 CAGCGTGTCACTACGAGCGGCGCGCGCGCGCGCGCTCTGCGGCGCTCTGCTGGG 2111
2101 CCGTACGATATCCACAGGCGCTGCGCACCTTCGCTGCTGCTGCGGCGCGCGAGCC 2160
2112 CCGTACGATATCCACAGGCGCTGCGCACCTTCGCTGCTGCTGCGGCGCGAGCC 2171
2161 GCCGCTGAGCTGTACTTTGTCAAGTGTGATGAGGCGCGGTACGACACCATCCCCA 2220
2172 GCCGCTGAGCTGTACTTTGTCAAGTGTGATGAGGCGCGGTACGACACCATCCCCA 2231
2221 GGACAGGTCACGAGGTCATCGCCAGCATCATCAACCCAGAGCATGCTACTGCTGG 2280
2232 GGACAGGTCACGAGGTCATCGCCAGCATCATCAACCCAGAGCATGCTACTGCTGG 2291
2281 TCGGTATGCGGTGTCAGAGAGCGCGCCATCGGACGTCGCCAAGGCTTCAAGAGCA 2340
2292 TCGGTATGCGGTGTCAGAGAGCGCGCCATCGGACGTCGCCAAGGCTTCAAGAGCA 2351
2341 GGTCTCTACCTTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGCTCCTCAGCA 2400
2352 GGTCTCTACCTTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGCTCCTCAGCA 2411
2401 GACACGCGCTGAGGATGCGGTGTCATCGACAGAGACTCCTCCCTGAATGAGGCGCAG 2460
2412 GACACGCGCTGAGGATGCGGTGTCATCGACAGAGACTCCTCCCTGAATGAGGCGCAG 2471
2461 CAGTGGCTCTTCGAGCTCTCTTCTAGCTTCTATGTCACACGCGGTGCGCATCAGGG 2520
2472 CAGTGGCTCTTCGAGCTCTCTTCTAGCTTCTATGTCACACGCGGTGCGCATCAGGG 2531
2521 CAAGTCTACGTCAGTCCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTG 2580
2532 CAAGTCTACGTCAGTCCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTG 2591
2581 CAGCTGTGCTACGCGGACATCGAGACAAGCTGTTTGGGGGATTCGCGGAGCGGCT 2640
2592 CAGCTGTGCTACGCGGACATCGAGACAAGCTGTTTGGGGGATTCGCGGAGCGGCT 2651
2641 GCTCTGCTGTTGTTGATGATTTCTTGTGTTGATGACCTCACCTCACCCAGCGAAGC 2700
2652 GCTCTGCTGTTGTTGATGATTTCTTGTGTTGATGACCTCACCTCACCCAGCGAAGC 2711
2701 CTTCTCAGGACCTGGTCCGAGGTGTCCTCTGAGTATGCTGCTGCTGAGTTCGGGAA 2760
2712 CTTCTCAGGACCTGGTCCGAGGTGTCCTCTGAGTATGCTGCTGCTGAGTTCGGGAA 2771
2761 GACAGTGTGACTTCCCTGTAGAGACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
2772 GACAGTGTGACTTCCCTGTAGAGACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2831
2821 GCCGCGCCACGGCTAT 2837
|||||
```

QY 2832 GCCGCGCCACGGCTAT 2848

RESULT 2 AF043739 8960 bp DNA SYN 05-APR-1998  
LOCUS  
DEFINITION Synthetic construct human telomerase catalytic subunit  
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete  
sequence.  
ACCESSION AF043739  
NID g3023054  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 8960)  
AUTHORS Counter, C.M., Meyerson, M., Eaton, E.N., Ellisen, L.W., Caddle, S.D.,  
Haber, D.A., and Weinberg, R.A.  
TITLE Telomerase activity is restored in human cells by ectopic  
expression of hTERT (hEST2); the catalytic subunit of telomerase  
Oncogene 16 (9), 1217-1222 (1998)  
JOURNAL  
MEDLINE 98187713  
REFERENCE 2 (bases 1 to 8960)  
AUTHORS Counter, C.M., Meyerson, M., Eaton, E.N., Ellisen, L.W., Caddle, S.D.,  
Haber, D., and Weinberg, R.A.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,  
Cambridge, MA 02142, USA  
FEATURES  
source  
1. 8960  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
order(1..1101,4606..8960)  
/note="mammalian expression vector; see GenBank Accession  
Number U47120"  
/organism="Cloning vector pCI-neo"  
/db\_xref="taxon:45772"  
1102..4542  
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.33"  
1102..4542  
/gene="hTERT/hEST2"  
1147..4605  
/note="telomerase/reverse transcriptase"  
/codon\_start=1  
/transl\_table=11  
/product="telomerase catalytic subunit"  
/db\_xref="PID:g3023055"  
/translation="MPRAPRCRAVRSLLRSHRYEVLPLATFVRLPGQWRLVQGGDP  
AAAFALVAQCLVCPVMDPPPAAPSPFQVSCLELVARVLQRCERKANLAFGEA  
LLDQAGGPPPEAFTTSVRSYLPNTVDALGSGAWGLLRVGDVLLVHLLALCALFY  
LVAPSCAYQVCGPLXOLGAOTAPPPHAGSPRRGLGCEANWHSYREAGVPLGLPA  
PGARRGCSASRLPLKPRRGAAPPEPTPVCGGSHWHPGRRGSDRGFCVVSFA  
RPAEATSLGALSGTRHSPSVGRHAGPPSTSPRPMDTFCPPVIAETKFLYS  
SGDQELRPSFLSSRLPTSLTGAARLVETIFLGSRPWMPGTPRRLPQRTWQMRPL  
FLELLGHAQCPYGLILKTHCPAAVTPAAGVCAAREKPGQSVAAPEEDTDRRLVQ  
LLRHSPPQVQVGFVRACLRLVPPGLWGRSHNRERFLRNTKFTISLKGAKLSLOEL  
TWKSVGCAWLRSPGCVPAAEHLREILAKFLHMLSVVVELLSFFVYET  
TFQNRLLFFRKSVWSKLSQIGIRHQLKRVQLRELSEAEVQREARPALTLRLRFT  
PKDGLRPIVMDYVGARTFRREKRAERTSRVKALFVNLNERRRPPGLGASVIG  
LDDTHRAWRTFVLRADQPPPELYFKVDVTGAYDTIPQDRLTEVASTIKPQNTYC  
VRRYAVVQKAAGHVRKAFKSHVSTLDLQAPYMRQFVAHQETSPLDADVIEQSSSL  
NEASSGLDFVLFREPMCHAVIRKGVSVQCGIPQGSITLSLCLSCVGDMEKLFAG  
IRRDGLLLRLVDDELLVTPHLTHAKTFLRLVIRGVPYEGCVVNLKTVNPNFVDEAL  
GGTAFVQMPAHGSLFWCGLLDDTTLTLEQSDYSIARTISIRASLTFNRGKAGNMR  
KLFGVLRKCHSLFDLVNSLTQVCTNIKLLQYRFRHACVLPQFPHQVWKNRT  
FFLRVISTASLCSYILKAKNAGSLGAKGAGPLPSEAVQWLCQHLKLTNRHT  
VPLGLSLRTAQTLQSLRKLPGTTLTALEAANPALPSDFKTLILDYPPVDYAGPYD  
4543..4605  
/note="engineered based on the epitope for the

anti-influenza virus hemagglutinin monoclonal antibody 12CA5; similar to part of the influenza A hemagglutinin gene sequence in GenBank Accession Number M29257" /organism="synthetic construct" /db\_xref="taxon:32630"

BASE COUNT	1886 a	2551 c	2484 g	2038 t	1 others
ORIGIN					

```
Query Match      99.2%; Score 2824; DB 31; Length 8950;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2825; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Query Match 99.2%; Score 2824; DB 31; Length 8960;

Best local similarity 99.56; P.E.A.M.O.: 0.00000;  
Matches 2826; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1101 CGTCTGCTGGCACGTGGGAGCCCTGGCCCGGCCACCCCGGATGCCGCGGCTCC 1160

QY 21 CGTCCTGCTGCGCACGTGGGAGCCCTGGCCCGGCCACCCCGCGATGCCGCGCGCTCC 80

Db 1161 CCGTGCCGAGCCGTCGGCTCCCTGCTGCGCAGCCACTACCGCAGGTGCTGCCGCTGGC 1220

QY 81 CCGCTGCCGAGCCGTGGCTCCCTGCTGCCGAGCCACTACCGGAGGTGCTGCCGTGGC 140

Db 1221 CACGTTGCGGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGACGCGGGGACCCGGC 1280

QY 141 CACGTCGTGCGCGCCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGGGGACCCGGC 200

Db 1281 GGCTTTCGGCGCTGGTGGCCCACTGGCTGGCTGGCGACGACGGCCGCC 1340

QY 201 GGCTTCCGGCGGCTGGTGGCCAGTGCCCTGGTGTGCGTGCCCTGGGACGCAGGCCGCC 260

Db 1341 CCCCCCGCCCCCTCCTTCCGCCAGGTGTCTCCTGAAGAGCTGGTGGCCCCGAGTGCT 1400

QY 261 C C C C G C C G C C C C C T C C T T C C G C C A G G T G T C C T G C C T G A A G G A G C T G G T G G C C C G A G T G C T 320

Db 1401 GCAGAGGCTGTGCGAGCGCGCGGCGAAGACGTGCTGGCCTTCGGCTTCGCGCTGCTGGA 1460

**QY**      321 GCAGAGGCTGTGCGAGCGGGCGGAACGTGCTGGCCTTCGGCTTCCGCCGTGCTGGA 380  
**QY**

Db 1461 CGGGGCCCGGGGGCCCCCGAGGCCTTCACCACCGCTGCGCAGCTACCTGCCCA 1520

QY 381 CGGGCCCCGGGGGGCCCCCGAGGCCCTCACCAACGCGTGCAGCTACCTGCCCA 440

Db 1521 CACGGTGACCGACGCACCTCGGGGGAGCGGGGGCTGCTGTTGCGCGCGTGGG 1580

QY 441 CACGGTGACCGACGCACTCGGGGGAGCGGGGGCTGCTGCTGCGCCGCGTGGG 500

Db 1581 CGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCTCTTTGTGCTGGTGGCTCCCAG 1640

QY 501 CGACGACGTGCTGGTTCACCTGCTGGCAGCTGCGGCTCTTTGTGCTGGTGGCTCCCAG 560

Db 1641 CTGCGCCTACCAAGTGTGCGGGCGCGCTGTACCAAGCTCGGGCTGCCACTCAGGCCCG 1700

QY 561 CTGGCCTACCAGGTGTGGGGCCGCCGCTGTACCAGCTGGCGCTGCCACTCAGGCCCG 620

Db 1701 GCCCCCGCCACACGCTAGTGGACCCCGAAGCGCTCTGGGATCGGAACGGGCTGGAACCA 1760

QY 621 GCCCCGCCACACGCTAGTGGACCCCGAAGCGCTGGGATGCGACGGCCCTGGAACCA 680

Db 1761 TAGCGTCAGGAGGCCGGGTCCCCCTGGGCTGCCAGCCCCGGGTGCGAGGAGCGCGG 1820

Qy 681 TAGCGTCAGGAGGCCGGGTCCCCCTGGGCTGCCAGCCCCGGGTGCGAGGAGCGCGG 740

Db 1821 GGGCAGTGCCAGCCGAAGTCTGCCGTGCCCAAGAGGCCAGGCGTGCCCTGA 1880

QY 741 GGGCAGTGCCAGCCGAAGTCTGCCGTGTGCCCAAGAGGCCAGGCGTGCCCTGA 800

Db 1881 GCCGAGCGGACGCCCGTTGGGCAGGGTCCTGGGCCACCCGGCAGGACGCGTGGACC 1940

QY 801 GCCGAGCGGACGCCCGTTGGCAGGGTCCTGGGCCACCCGGCAGGACGGTGGACC 860

Db 1941 GAGTGACCGTGGTTTCTGTGTGGTGTCACTGCCAGACCCGCCGAGAACCTCTTT 2000

QY 861 GAGTGACCGTGTTCCTGTGGTGTCACCTGCCAGACCCGCCGAGAACCTCTTT 920

Db 2001 GGAGGGTGGGCTCTCTGGCAGCGGCGCACTCCACCCATCCGTGGCCGCGCACCAAGC 2060



|||||  
QY 262 CCGCGGCCCCCTCTTCGCGCAGGTGCTCTGCTGAAGAGAGCTGGTGGCCCCAGTGTG 321  
Db 312 CAGAGGCTGTGCAGCGCGCGCGAGAAAGTGTCTGGCTTCGGCTTCGCGTGTCTGGAC 371  
QY 322 CAGAGGCTGTGCAGCGCGCGCGAGAAAGTGTCTGGCTTCGGCTTCGCGTGTCTGGAC 381  
Db 372 GGGGCCCCGGGGCCCCCCCCAGGCTTCACACACAGAGTGTGGAGTACCTGCCCCAAC 431  
QY 382 GGGGCCCCGGGGCCCCCCCCAGGCTTCACACACAGAGTGTGGAGTACCTGCCCCAAC 441  
Db 432 ACGGTGACCGACGACTGTGGGGGAGCGGGGTGGGGCTGCTGTGGCGCGGTGGGC 491  
QY 442 ACGGTGACCGACGACTGTGGGGGAGCGGGGTGGGGCTGCTGTGGCGCGGTGGGC 501  
Db 492 GACGAGTGTGCTGTACCTGTGTCACGCTGCGGCTCTTTGTGTGTGTGTGCTCCACG 551  
QY 502 GACGAGTGTGCTGTACCTGTGTCACGCTGCGGCTCTTTGTGTGTGTGTGCTCCACG 561  
Db 552 TGGGCTTACAGGTGTGGGGCGCGCGCTGTACAGCTCGGCGCTGCGACTCAGGCGCGG 611  
QY 562 TGGGCTTACAGGTGTGGGGCGCGCGCTGTACAGCTCGGCGCTGCGACTCAGGCGCGG 621  
Db 612 CCCC CGCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGGAAGCGGCTTGAACCAT 671  
QY 622 CCCC CGCACACGCTAGTGGACCCCGAAGCGCTCTGGGATGGAAGCGGCTTGAACCAT 681  
Db 672 ACGGTCAGGAGCGCGGCTCCCTTGGGCTTCCAGGCCCCGGGTGCGAGAGGCGCGG 731  
QY 682 ACGGTCAGGAGCGCGGCTCCCTTGGGCTTCCAGGCCCCGGGTGCGAGAGGCGCGG 741  
Db 732 GCGAGTGCAGCGAAGTGTGCGCTTGTCCCAAGAGCCCGAGCGGTGGCGCTGCCCTGAG 791  
QY 742 GCGAGTGCAGCGAAGTGTGCGCTTGTCCCAAGAGCCCGAGCGGTGGCGCTGCCCTGAG 801  
Db 792 CCGGAGCGAGCCCGTGTGGGAGGGGTCTGCGGCCACCGCGGAGAGCGGTGGACCG 851  
QY 802 CCGGAGCGAGCCCGTGTGGGAGGGGTCTGCGGCCACCGCGGAGAGCGGTGGACCG 861  
Db 852 AGTGACCGTGTGTTCTGTGTGTGTGTCACCTGCGACACCGCGCGAAGAGCCACCTTTTG 911  
QY 862 AGTGACCGTGTGTTCTGTGTGTGTGTCACCTGCGACACCGCGCGAAGAGCCACCTTTTG 921  
Db 912 GAGGTTGGCTCTCTGGACGCGGCACCTCCACCATCGTGGGCGCGCAGACACCGG 971  
QY 922 GAGGTTGGCTCTCTGGACGCGGCACCTCCACCATCGTGGGCGCGCAGACACCGG 981  
Db 972 GGGCCCCATCCACATCGGGGCCACACGCTCCCTGGGACACGCTTGTCCCGCGGTGAC 1031  
QY 982 GGGCCCCATCCACATCGGGGCCACACGCTCCCTGGGACACGCTTGTCCCGCGGTGAC 1041  
Db 1032 GCCGAGACCAAGCACTTCTCTACTCTCTCAGCGACAGGAGACGCTGGCGCCCTCTTC 1091  
QY 1042 GCCGAGACCAAGCACTTCTCTACTCTCTCAGCGACAGGAGACGCTGGCGCCCTCTTC 1101  
Db 1092 CTACTCAGGCTCTCTGAGGCCAGCTGACTGGCGTTCGAGGCTCGTGGAGACCACTTTT 1151  
QY 1102 CTACTCAGGCTCTCTGAGGCCAGCTGACTGGCGTTCGAGGCTCGTGGAGACCACTTTT 1161  
Db 1152 CTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGGCTGCCCGAGGCG 1211  
QY 1162 CTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGGCTGCCCGAGGCG 1221  
Db 1212 TACTGGCAATCGGCCCCCTGTTCTTGGAGCTGCTTGGGAACACAGCGAGTGGCCCTTAC 1271  
QY 1222 TACTGGCAATCGGCCCCCTGTTCTTGGAGCTGCTTGGGAACACAGCGAGTGGCCCTTAC 1281  
Db 1272 GGGGTGCTCTCAAGACGACTGCCCCGTGCGAGTGTGGGTCACCCCAGACGCGGTGTC 1331  
QY 1282 GGGGTGCTCTCAAGACGACTGCCCCGTGCGAGTGTGGGTCACCCCAGACGCGGTGTC 1341  
Db 1332 TGTGCGCGGAGAACCCCGAGGGCTCTGTGGCGGCCCGCGAGGAGGAGACACACGCC 1391  
|||||

QY 1342 TGTGCCCGGGAAGAACCCAGGGCTCTGTGGCGGCCCGGAGGAGGACACAGACCCC 1401  
Db 1392 CGTCCCTGTGTGACGCTGCTCCGCGACACACAGACGCCCTTGGCAGGTGTACGGGTTCGTG 1451  
QY 1402 CGTCCCTGTGTGACGCTGCTCCGCGACACACAGACGCCCTTGGCAGGTGTACGGGTTCGTG 1461  
Db 1452 CGGGCTGCTGTGGCGGGCTGGTGCCTCCAGGCTCTGGGGCTCCAGGACCAAGAACGC 1511  
QY 1462 CGGGCTGCTGTGGCGGGCTGGTGCCTCCAGGCTCTGGGGCTCCAGGACCAAGAACGC 1521  
Db 1512 CGCTTCCTCAGGAACACCAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTTCGCTG 1571  
QY 1522 CGCTTCCTCAGGAACACCAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTTCGCTG 1581  
Db 1572 CAGGAGTGTGCTGGAAGATGAGGTGCGGGGCTGCGCTTGGCTGCGCAGAGCCAGGG 1631  
QY 1582 CAGGAGTGTGCTGGAAGATGAGGTGCGGGGCTGCGCTTGGCTGCGCAGAGCCAGGG 1641  
Db 1632 GTTGGCTGTGTTCCGGCGCAGAGACGCTGTGCTGTGAGGAGATCTTGGCCAAAGTTCCTG 1691  
QY 1642 GTTGGCTGTGTTCCGGCGCAGAGACGCTGTGCTGTGAGGAGATCTTGGCCAAAGTTCCTG 1701  
Db 1692 CACTGGCTGTAGTGTGTACGTGCTGAGCTGCTCAGGCTCTTTCTTTATGTACCGAG 1751  
QY 1702 CACTGGCTGTAGTGTGTACGTGCTGAGCTGCTCAGGCTCTTTCTTTATGTACCGAG 1761  
Db 1752 ACCAGCTTTCAAAAGAACAGGCTCTTTTCTTACCGGAAGAGTGTCTGGAGCAAGTTGCAA 1811  
QY 1762 ACCAGCTTTCAAAAGAACAGGCTCTTTTCTTACCGGAAGAGTGTCTGGAGCAAGTTGCAA 1821  
Db 1812 AGCATTTGAATCAGACAGCACTTGAAGAGGTGAGCTGTGCGGAGCTGTCTGGAGAGCAGAG 1871  
QY 1822 AGCATTTGAATCAGACAGCACTTGAAGAGGTGAGCTGTGCGGAGCTGTCTGGAGAGCAGAG 1881  
Db 1872 GTCAGGAGCATCGGAGAGCCAGCCCGCTGCTGAGCTCCAGCTCCGCTTCATCCCC 1931  
QY 1882 GTCAGGAGCATCGGAGAGCCAGCCCGCTGCTGAGCTCCAGCTCCGCTTCATCCCC 1941  
Db 1932 AAGCTCAGCGGCTGCGGCGGATTTGTAAACATGGACTACGTCGTGGGAGCCAGAACGTTTC 1991  
QY 1942 AAGCTCAGCGGCTGCGGCGGATTTGTAAACATGGACTACGTCGTGGGAGCCAGAACGTTTC 2001  
Db 1992 CGCAGAGAAAGAGGGCGGAGCGTCTCAGCTCAGAGGTGAAGGCACTGTTTCAGCGTGTGTC 2051  
QY 2002 CGCAGAGAAAGAGGGCGGAGCGTCTCAGCTCAGAGGTGAAGGCACTGTTTCAGCGTGTGTC 2061  
Db 2052 AACTACAGGCGGCGCGCGCCCGCTCTGCGGCGCTCTGCTGGGCTTGAGCAT 2111  
QY 2062 AACTACAGGCGGCGCGCGCCCGCTCTGCGGCGCTCTGCTGGGCTTGAGCAT 2121  
Db 2112 ATCCACAGGGCTGGCGCACCTTCTGCTGTGCTGGGGCCCGAGGACCCCGCCTGAG 2171  
QY 2122 ATCCACAGGGCTGGCGCACCTTCTGCTGTGCTGGGGCCCGAGGACCCCGCCTGAG 2181  
Db 2172 CTGTACTTTTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCGCAGGACAGGCTC 2231  
QY 2182 CTGTACTTTTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCGCAGGACAGGCTC 2241  
Db 2232 ACGGAGTCTATCGCCGACATCATCAAAACCCAGAACAGTACTGCTGCTGGTATGTC 2291  
QY 2242 ACGGAGTCTATCGCCGACATCATCAAAACCCAGAACAGTACTGCTGCTGGTATGTC 2301  
Db 2292 GTGTTCAGAGAGGGCGGCCATGGCAGCTCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2351  
QY 2302 GTGTTCAGAGAGGGCGGCCATGGCAGCTCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2361  
Db 2352 TTGACAGACTTCCAGCGGTACATCGGAGTTCGTGGTCACTTGCAGGAGACAGGCCG 2411  
QY 2362 TTGACAGACTTCCAGCGGTACATCGGAGTTCGTGGTCACTTGCAGGAGACAGGCCG 2421  
Db 2412 CTGAGGAGTCCGCTGCTATCGGAGCAGAGCTCTCTGATGAGGCGAGCAGTGGGCTC 2471  
QY 2422 CTGAGGAGTCCGCTGCTATCGGAGCAGAGCTCTCTGATGAGGCGAGCAGTGGGCTC 2481

Db	2472	TTCCAGCGTCTTCTTACGCTTCATGTGCCACCACGCGGTGGCATCAGGGGCAAGTCCTAC	2531
Qy	2482	TTCCAGCGTCTTCTTACGCTTCATGTGCCACCACGCGGTGGCATCAGGGGCAAGTCCTAC	2541
Db	2532	GTCCAGTGCACGGGATCCCGCAGGGCTCCATCTCTCCAGCTGCTCTGCAAGCTCTGC	2591
Qy	2542	GTCCAGTGCACGGGATCCCGCAGGGCTCCATCTCTCCAGCTGCTCTGCAAGCTCTGC	2601
Db	2592	TACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGACGGGCTGCTCCTCGT	2651
Qy	2602	TACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGGACGGGCTGCTCCTCGT	2661
Db	2652	TTGGTGGATGATTTCTTGTGTGTGACACCTCACTCAACCCACGCGGAAACCTTCTCAGG	2711
Qy	2662	TTGGTGGATGATTTCTTGTGTGTGACACCTCACTCAACCCACGCGGAAACCTTCTCAGG	2721
Db	2712	ACCTTGGTCCGAGGTCTCCTCAGTATGCTGCGTGGTGAACCTTGGGAAGACAGTGGTG	2771
Qy	2722	ACCTTGGTCCGAGGTCTCCTCAGTATGCTGCGTGGTGAACCTTGGGAAGACAGTGGTG	2781
Db	2772	AACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGGCTTTTGTTCAGATGCGCGGCCAC	2831
Qy	2782	AACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGGCTTTTGTTCAGATGCGCGGCCAC	2841
Db	2832	GGCCTAT 2838	
Qy	2842	GGCCTAT 2848	
RESULT	4		
LOCUS	AF051911	3426 bp mRNA	ROD
DEFINITION	Mus musculus telomerase reverse transcriptase mRNA, complete cds.		02-APR-1998
ACCESSION	AF051911		
NID	93005591		
KEYWORDS	house mouse.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 3426)		
AUTHORS	Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.		
TITLE	Expression of mouse telomerase reverse transcriptase during Oncogene (1998) In press		
JOURNAL	2 (bases 1 to 3426)		
REFERENCE	Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-MAR-1998) Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA		
FEATURES	Location/Qualifiers		
source	1..3426		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="13"		
	/map="MGD offset position 40; 40 cm from centromere"		
CDS	30..3398		
	/note="mTEXT"		
	/codon_start=1		
	/product="telomerase reverse transcriptase"		
	/db_xref="PID:g3005592"		
	/translation="MTRAPRCPAVRSLRSRYREVWPLATFVRRLGPEGRLVQPCGD KTYRLVAQCLVCHWGSOPPADLSFHOVSSLKELVAVYVQRLCERNERNVLVAGF LLNEARGPPMAFVSRSYLPNTVETLRVSGAWMLLSRVGDDLLVLLAHACALY LVPPSCAYVCGSPYLQICATDITWPSVASYRPTREVGNSTNLFQLOIKSSSQ APKPLALPSRGTKHLISUTSVPSAKKACYPVPRVEEGPHQVLPFTSGSKSWPS ARSPVPAEDKLSKQKVDLSLGSVCCKHKPSSILSPQRONAFOLRPFIEFR FLYSGDGOERLNTSFLLSNLQNLGTGARLVEIFIGSRPRTSGPLCRTLRLSPW QMRPLFOQLLYNHAEQVYVLLRSHCFRTANQVTDALNTSPHMLDLLRLRHSPW VYGFUACCLVYNSASLGMTRHNERFFKNLKFISLUGKYKLSLQELQMLWKKRVEDC WRSSPGRDQVPAAPHLRERILATLFLWLDLTVQVLLRSFFYITESTFTQNRLEF RLRSVMSKQSVLQVHLRERLSEQVEVHHQDTWLAMPICRLARFIPKPNGLREI		

BASE COUNT		746 a	979 c	907 g	794 t
ORIGIN					
Query Match		29.0%	Score 825;	DB 28;	Length 3426;
Best Local Similarity		67.1%;	Pred. No. 0.00e+00;		
Matches 1901; Conservative		0;	Mismatches 838;	Indels 93;	Gaps 5;
Db	14	CCCGGCGCTTGAGCACAATGACCCGCGTCTCTCGTTGGCCCGCGTGGCGCTCTCTGCTGCG 73			
QY	51	CCCGGCCACCCCGGATGCGCGCGTCCCGCTGCGAGCCGTGGCTCCCTGCTGCG 110			
Db	74	CAGCGGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGCGGCGCCTGGGGCCCGAGGG 133			
QY	111	CAGCCACTACCGGAGGTGTGCGCGTGGCCACGTTCTGTGCGGCGCCTGGSGCCCCAGGG 170			
Db	134	CAGCGGCTGTGCAACCCGGGGACCCGAGATCTACCGCACTTTGGTTGGCCCAATGCCT 193			
QY	171	CTGCGCGCTGTGTGAGGCGGGGACCGCGCGCTTTCCGCGCGTGTGTGGCCAGTGCCT 230			
Db	194	AGTGTGATGCACGTGGGGCTACAGCCTCCACCTCGCGGACCTTTCCCTTCCACGAGGTGC 253			
QY	231	GGTGTGGGTGCGCTTGGGACGACGGCGCGCCCCCGCGCGCCCCCTCTTCCGCCAGGTGTC 290			
Db	254	ATCCCTGAAGAGCTGTGTGGCGAGGTTGTGCAGAGACTCTGCGAGGCGCAACGAGAGAAA 313			
QY	291	CTGCCTGAGGAGCTGTGTGCCCGAGTGTGTGACAGGCTGTGCGAGGCGCGCGCGGAGAA 350			
Db	314	CGTCTGGCTTTTGGCTTTTGAGCTGTCTTAACGAGGCGCAGAGCGGGCGCTCCCATGGCCTT 373			
QY	351	CGTCTGGCTTTCGGCTTTCGGCTGCTGGACGGGGCGCGCGGGGCGCCCCCGGAGGCCCTT 410			
Db	374	CACGTAGTAGCGGTGAGTACTTTGCCCAACACTGTTATTGAGACCTTCGCTGTCTAGTGG 433			
QY	411	CACCACGAGGTGGCAGCTACTTCCGCGCAACACAGGTGACCGACGACTGTGCGGGGAGCGG 470			
Db	434	TGCATGATGCTACTGTTAGCCGAGTGGGGACGACCTGCTGCTCTACCTGCTGGCACAC 493			
QY	471	GGCTGGGGGCTGTGCTGCGCGCGTGGGCGACGACGTGCTGTTTCACTGCTGGCAGG 530			
Db	494	CTGTGCTCTTTATCTTCTGTGTGCCCGCCAGCTGTGCCCTACCAAGTGTGTGGGTCTCCCT 553			
QY	531	CTGCGCGCTCTTTGTGTGTGGTGTCCAGTGTGGCCTACCAAGTGTGTGCGGCGCCGCT 590			
Db	554	GTACCAATTTGTCCACCGGATATCTGGCCCTCTGTGTCCGCTAGTTACAGGCCAC 613			
QY	591	GTACCACTCGGCGCTGCCACTGAGGCCGCGCCCCCGCCACACGCTAGTGGACCCCGAAG 650			
Db	614	CCGACCGTGGCGAGGAATTTCACTAACCTTTAGTTCTTTACAACGATCAAGACGATG 673			
QY	651	CGTCTGGGATGCG-----AAGGGCGCTGGAAACCATAG 683			
Db	674	TCGCGAGGAGCACCAGAACCCCTTGGCCTTGCCTATCTGAGGTACAAGAGGCATCTGAG 733			
QY	684	CGTCAGGAGGCCGGGTGCCCTTGGGCTTGCAGCCCCGGGTGCGAGGAGCGCGGGGG 743			
Db	734	TCTCACCAGTACAAGTGTGCCTTCAGCTTAAGAGGCCAGAGTGCATCTGTCGCGAGAGT 793			
QY	744	CAGTGCCAGCGGAAGTCTGCGGTTTGGCCAAAGAGGCCACGCGTGTGCGCTGAGCC 803			
Db	794	GGAGGAGGCCCCACAGGAGGTGCTACCAACCCCATCAGGCAATCATGGGTGCCAAG 853			
QY	804	GGAGCGACGCCCGTGGCGAGGGGTCTGTGGGCGCCACCCGGCGAGGACGCGTGACCGAG 863			
Db	854	TCCTGTCTGGTCCCGGAGTGGCTACTGCAAGAGAAAGATTGTCTCTTAAGAGAAAGGT 913			





[illegible]

Db	421	TTGAGCGAGTGGGCGACGACCTGCTGTGGTCTACCTGCTGGCACACTGTCCTTTATCTT	480
Qy	487	CTGGCGCGGTGGGCGACGCTGCTGGTTTACCTGCTGGCAGCGTGCGCGCTCTTTGTG	546
Db	481	CTGGTGGCCCCAGCTGTGCCTACCAAGTGTGTGGGTCTCCCTGTACCAAAATTTGTGCC	540
Qy	547	CTGGTGGCTCCCAAGTGTGCCTTACCAAGTGTGTGGGGCGCGCGTGTACCAAGCTCGGGCGCT	606
Db	541	ACCACGGATATCTGGCGCTCTGTGTCGGCTAGTCTACAGGCCACCCGACCGCTGGGCAGG	600
Qy	607	GCACCTACGGCGGGCCCCGCCACACGCTAGTGACCCCGAAGCGTCTGGATGCG--	664
Db	601	AATTTCACTAACCTTAGTCTTTACAACAAGATCAAGACGATGTGCGCAGGAACACCG	660
Qy	665	-----AACGGGCTTGAACCATAGCTACGGGAGCGCGG	699
Db	661	AAACCCCTGGCCTTGCCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACCAAGTACAAAGT	720
Qy	700	GTCCCTCTGGGCTGCGACCCCGGGTGGAGAGGCGCGGGGCGAGTGCACAGCCGAAGT	759
Db	721	GTGCTTCAAGTAAAGAGGCGAGATGCTATCTGTCCGAGAGTGGAGAGGAGGCCAC	780
Qy	760	CTGCGTGTGCCAAGAGGCCAGGGCTGGCGTGCCTCTGAGCGGAGCGAGCCCGT	819
Db	781	AGGACAGTGTCTACAAACCCCATCAGSCAAATCATGGGTGCCAAGTCTCTGCTGCCCTCC	840
Qy	820	GGGACGGGTCTTGGCCCCACCGGAGGCGGTGGACCGAGTGACCGTGGTTCTGT	879
Db	841	GAGTGGCTACTGCAGAGAAAGATTGTCTTAAAGAAAGGTGTCTGACCTGAGTCTC	900
Qy	880	GTGGTGTCACTGCCAGACCCCGCAAGAACCACTCTTTGGAGGGTGCCTCTCTGGC	939
Db	901	TCGGG-----TCGGTGTCTCTAAACAACAGCCAGCTCCACACTCTCTGCT	948
Qy	940	ACGGGCCACTCCACCCCTCCGTGGGCGCCAGCACACAGGGGGCCCCCATCCACATCG	999
Db	949	TCACCAACCCCGCCAAATGCCTTTCAGCTCAGGCCATTTATTGAGACCAGACATTTCTT	1008
Qy	1000	CGGCCACACGTCCTGGGACAGCCTTGTCCCCGGTGTACGCCGAGACCAAGCACTC	1059
Db	1009	TACTCCAGGGGAGATGGCAAGAGCGCTTAAACCCCTCATTTCTTACTCAGCAACCTCAG	1068
Qy	1060	CTCTACTCTCTCAGGCACACAGGACGAGCTGCGGGCCCTCTCTCTACTCAGCTCTGTAGG	1119
Db	1069	CCTAACCTGACTGGGCGCAGGAGCTGGTGGAGATCACTTTCTGGGCTCAAGGCTTAG	1128
Qy	1120	CCAGCCTGACTGGCGCTCGGAGGCTGTGGAGACCATCTTCTGGGTTCCAGGCCCTGG	1179
Db	1129	ACATCAGGACCACTCTGCAGGACACACCCCTCTATCGCGCTGATACTGGCAGATCGGGCC	1188
Qy	1180	ATGCCAGGACTCCCGCAGGTTGCCCGCCTGCCCCAGCGCTACTGCAATCGGGGCC	1239
Db	1189	CTGTTCACAAGCTGTGTGTGAACCATGCAGAGTGCCCAATATGTCTAGACTCCTCAGGTCA	1248
Qy	1240	CTGTCTTGAGCTGTGTGGGAACCAACGCGCAGTGCCTCTACGGGTCTCTCTCAAGACG	1299
Db	1249	CATTGCAAGTTTTCACACACCAACCAAGGTGCAGATGCCTT-----	1292
Qy	1300	CAGTCCCGCTGCGAGCTCGGTCTACCCAGACGCGGTGTCTGTGCCCGGAGAACCC	1359
Db	1293	-----GAACACAGCCCAACCGACTCATGGATTG	1323
Qy	1360	CAGGGCTCTGTGGGGCCCCCGNGAGGAGGACACAGACCCCGCTGCCTGGTGCAGCTG	1419
Db	1324	CTCGCTGCACAGCAAGTCCCTGGCAGGTATATGTTTTCTTGGGGCTGTCTCTGCAAG	1383
Qy	1420	CTCGCCAGCACAGCACCCCTGGCAGGTGTACGGTCTGTGCGGGCTGCCTCGCGCG	1479
Db	1384	GTGGTGTCTGTAGTCTGTGGGTACCAGGCACATGAGCGCCGTCTTTAAGAACTTA	1443
Qy	1480	CTGTGTCCTTGGGCTCTGGGCTCCAGGCACACAGAACCGGCTTCTCAGGAACACC	1539

```
Db 1444 AAGAGTTTCATCTCGTTGGGAAATACGGCAAGCTATCACTCCAGGAACCTGATGTGGAG 1503
|||
Qy 1540 AAGAGTTTCATCTCGTTGGGAAAGATGCCAAGCTCTCGTGCAGGAGCTGACCTGGAG 1599
|||
Db 1504 ATGAAGTAGAGGATTCGCACTGGCTCGGACAGACGCCGGGAAAGACCGTGTCCCGGCT 1563
|||
Qy 1600 ATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGGAGCCAGGGTTGGTGTGTTCGGGCC 1659
|||
Db 1564 GCAGAGCACCGTCTGAGGAGGAGGATCTGGCTAGTTCCTGCTGCTGATGAGACACA 1623
|||
Qy 1660 GCAGAGCACCGTCTGCTGAGGAGATCTGGCCAAAGTTTCTGCACTGCTGATGAGTGTG 1719
|||
Db 1624 TACGTGGTACAGCTGCTTAGTTCATCTTTACATCACAGAGAGCAATCCAGAGAAC 1693
|||
Qy 1720 TACGTGCTGAGCTGCTCAGGTCTTCTTTATGTACGAGGACACCGTTTCAAAAGAAC 1779
|||
Db 1684 AGGCTCTTCTTCTACCGTAAGAGTGTGGAGCAAGCTGCAGAGCATGGAGTCAGGCAA 1743
|||
Qy 1780 AGGCTCTTCTTCTACCGTAAGAGTGTGGAGCAAGCTGCAGAGCATGGAGTCAGGCAA 1839
|||
Db 1744 CACCTTGAGAGAGTGGCTACGGGAGCTGTCAACAAGAGAGGTTCAGGCAATCACAGGAC 1803
|||
Qy 1840 CACTTGAAGAGGTGCAGCTCGGAGCTGTGGAAGCAGAGGTTCAGGCAATCACAGGAC 1899
|||
Db 1804 ACCTGGCTAGCATGCCATCTGCAGACTGCGTTCATCCCAAGCCCAAGCGGCTGGG 1863
|||
Qy 1900 GCCAGGCCGCGCTGCTGACCTCCAGACTCCGCTTCATCCCAAGCCCTGACGGCTGGG 1959
|||
Db 1864 CCCATTGTGAACATGAGTTATAGCATGGGTACCAAGAGCTTTGGCAGAGGAAGCAGGCC 1923
|||
Qy 1960 CCGATTGTGAACATGAGTTATAGCATGGGTACCAAGAGCTTTGGCAGAGGAAGCAGGCC 2019
|||
Db 1924 CAGCATTTCCACCGCTCTCAAGACTCTTCTGACATGCTCAACTATGAGCGACAAAA 1983
|||
Qy 2020 GAGCGTCTCACTCGAGGTTGAAGCACTGTTGAGCGTGTCAACTGAGCGGCGCGG 2079
|||
Db 1984 CATCTCACCTTATGGGCTCTCTGACTGGGTATGATGACATCTACAGACCTGGCGG 2043
|||
Qy 2080 CGCCCGCGGCTCTGGGCGGCTCTGTGCTGGGCGTGGAGATATCCACAGGCGCTGGCG 2139
|||
Db 2044 GCCTTTGTGCTGCGTGTGCTGACACAGACACCCAGGATGTACTTTGTTAAGGCA 2103
|||
Qy 2140 ACTTGTGCTGCGTGTGCGGCGGCGGAGACCCGCGCTGAGCTGTACTTTGTCAAGTG 2199
|||
Db 2104 GATGTACCGGGGCTATGATGCCATCCCGAGGCTAAGCTGGTGGAGTTGTTGCCAAT 2163
|||
Qy 2200 GATGTACCGGGGCTATGATGCCATCCCGAGGCTAAGCTGGTGGAGTTGTTGCCAAT 2259
|||
Db 2164 ATGATCAGGCACTCGGAGAGCAGCTACTGTATCCGCCAGTATGAGTGGTCCGGAGAGAT 2223
|||
Qy 2260 ATCATCAACACCC---AGAACACGTACTGCGTGTGCGTATGCGGTGGTCCAGAGGCC 2316
|||
Db 2224 AGCCAAGGCCAAGTCCCAAGTCTTTAGGAGACAGAGTTCACACCTCTCTGACCTCCAG 2283
|||
Qy 2317 GCCATGGGACGTCGCGAAGGCTTCAAGAGCACGCTCTACTCTGACAGACCTCCAG 2376
|||
Db 2284 CCATACATGGGCCATGCTTCAAGCATCTGCAGGATTCAGATGCCAGTGCCTACTGAGGAAC 2343
|||
Qy 2377 CCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAGAC-----CAGCCCGCTGAGGAT 2430
|||
Db 2344 TCCGTTGTATCAGCAGAGCATCTATGATGAGAGCAGCAGCAGCTGTTTGAAGTTC 2403
|||
Qy 2431 GCCGTCGTCATCAGCAGAGATCTCTCCCTGATGAGGCCAGCAGTGGCCCTTTCGAGGTC 2490
|||
Db 2404 TTCCTGACTTCTCTGCGTCACAGTGTCTGTAAGATTTGGTGTGACAGTGTCTATACAGTGC 2463
|||
Qy 2491 TTCTACGCTTATGTGCCACACCGCTGCGCATCAGGGCAAGTCTCTAGTCCAGTGC 2550
|||
Db 2464 CAGGGCATCCCGGCTCCAGGCTATCCACCTGCTCTGAGTCTGTGTTTCGAGAGAC 2523
|||
Qy 2551 CAGGGGATCCCGGCTCCAGGCTATCCCTCTCCACGCTGCTCTGAGCCTGTGCTACGCGGAC 2610
|||
Db 2524 ATGAGAACACAGCTGTTTGTGTGAGGTGCAGCGGGATGGGTTGCTTTTACGTTTTTGTAT 2583
|||
Qy 2611 ATGAGAACAGCTGTTTGTGGGGATTCGCGGAGCGGCTGCTCTCGCTTTGGTGGAT 2670
|||
Db 2584 GACTTTCTGTTGTTGAGCCTCACTTGGACCAAGCAAAAAACCTTCTCTCAGCACCTGGTC 2643
|||
Qy 2671 GATTTCTGTTGTTGAGCCTCACTTGGACCAAGCAAAAAACCTTCTCTCAGCACCTGGTC 2730
|||
Db 2644 CATGGCGCTTCTGATGATGGTGCATGATAAACTTGCAGAGAGAGTGGTGAACCTCCCT 2703
|||
Qy 2731 CGAGGTCTCCCTGATGATGCTGCTGTAACCTTGGGAGAGACAGTGGTGAACCTCCCT 2790
|||
Db 2704 GTGAGCCTGGTACCCCTGGTGGTGCAGCTTCCATACAGCTGCTGCTCACTGCT 2759
|||
Qy 2791 GTAGAGAGAGCGGCTGGTGGTGCAGCTTGTGTTTTCAGATGCCGCGCCAGCGCT 2846
|||
RESULT 6
LOCUS AF029235 201 bp mRNA ROD 12-NOV-1997
DEFINITION Mus musculus telomerase catalytic subunit mRNA, partial cds.
ACCESSION AF029235
NID 92605902
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
REFERENCE 1 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Partial sequence of Mus musculus telomerase catalytic subunit
homolog
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.L.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1997) Biochemistry, St Jude Children's Research
Hospital, 332 North Lauderdale, Memphis, TN 38105, USA
FEATURES
source
1..201
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="32D.3"
<1..>201
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g2605903"
/translation="FFVTFSTFKNRLFFKYSVMSKLSQISGVRLRLRLSLSQ
EVRHHDWFLAMPICRLRFIPK"
BASE COUNT 51 a 54 c 55 g 41 t
ORIGIN
Query Match 4.5%; Score 129; DB 28; Length 201;
Best Local Similarity 82.1%; Pred. No. 1.63e-71;
Matches 165; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Db 1 TTCTTTTATGTCAGGAGACACATTCAGAGAACAGGCTCTTCTTCTACCTAAGAT 60
Qy 1744 TTCTTTTATGTCAGGAGACACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAT 1803
|||
Db 61 GTGTGAGCAGCTGCAGAGCATTTGGATGAGTGCAGCAACACCTTGAGAGTGCAGGCTACGG 120
Qy 1804 GTGTGAGCAGTGTGCAAGAGCATTTGAAATCAGACAGCATTTGAAGAGGTGCAGCTGCGG 1863
|||
Db 121 GAGTGTCAACAAGAGGAGTGCAGGATCACCAGGACACCTGAGCTAGGCTGCGCATGCTGC 180
Qy 1864 GAGTGTCAACAAGAGGAGTGCAGGATCACCAGGACACCTGAGCTAGGCTGCGCATGCTGC 1923
|||
Db 181 AGACTCGCTTTCATCCCAAG 201
Qy 1924 AGACTCGCTTTCATCCCAAG 1944
|||
RESULT 7
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
```



<hr/>						
DEFINITION	Sequence 14 from patent US 5670367.					
ACCESSION	I66494					
NID	92724471					
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 7218)					
AUTHORS	Dorner,F., Scheifflinger,F. and Falkner,F.Gunter.					
TITLE	Recombinant fowlpox virus					
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;					
FEATURES	Location/Qualifiers					
source	1..7218					
BASE COUNT	1944 a 1491 c 1486 g	1929 t	368 others			
ORIGIN	/organism="unknown"					
<hr/>						
Query Match	2.4%; Score 67; DB 21; Length 7218;					
Best Local Similarity	2.4%; Pred. No. 1.13e-25;					
Matches	9; Conservative 211; Mismatches 153; Indels 0; Gaps 0;					
<hr/>						
Db	1055 GGAGCTTGCATY	YYY	1114			
	:::	: : : : : ::::: : : : : : : : : : : : : : : :	:	:	:	:
Qy	921 GGAGGTGCGCTCTCTGGCAGCGCCACTCCCACCATCGTG	GGCGCGCCAGCACGCGC	980			
<hr/>						
Db	1115 YYY	YYY	1174			
	::::: :	:	:	:	:	:
Qy	981 GGGCCCCCATCCACATCGGGGCCACCACTGCCCTGGGACACGCTGT	GCCCCGGTGTA	1040			
<hr/>						
Db	1175 YYY	YYY	1234			
	::: :	:	:	:	:	:
Qy	1041 CGCCGAGACAAGCACTTCCTCTACTCTCAGGACAAGGACGAGT	CGCGCCCTCT	1100			
<hr/>						
Db	1235 YYY	YYY	1294			
	::: :	:	:	:	:	:
Qy	1101 CCTACTAGCTCTCTGAGCCAGCCTGACTGGCGCTCGGAGGCTCGT	GGAGACAATCT	1160			
<hr/>						
Db	1295 YYY	YYY	1354			
	::: :	:	:	:	:	:
Qy	1161 TCTGGTTCCAGGCGCTGGATGCCAGGACTTCCCGCAGGTTG	CCCGCGCTGCCCGCAGCG	1220			
<hr/>						
Db	1355 YYY	YYY	1414			
	::: :	:	:	:	:	:
Qy	1221 CTACTGGCAATCGGCCCTCTTCTTGAGACTGCTTGGGAACCA	CAGCGAGTGCCCCCTA	1280			
<hr/>						
.Db	1415 YYY	YYYYYYYYYYYYY	1427			
	: : : : : : : : :	:	:	:	:	:
Qy	1281 CGGGTGCTCTC	1293				
<hr/>						
RESULT	8					
LOCUS	I66494	7218 bp	DNA	PAT	23-DEC-1997	
DEFINITION	Sequence 14 from patent US 5670367.					
ACCESSION	I66494					
NID	92724471					
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 7218)					
AUTHORS	Dorner,F., Scheifflinger,F. and Falkner,F.Gunter.					
TITLE	Recombinant fowlpox virus					
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;					
FEATURES	Location/Qualifiers					
source	1..7218					
BASE COUNT	1944 a 1491 c 1486 g	1929 t	368 others			
ORIGIN	/organism="unknown"					
<hr/>						
Query Match	1.7%; Score 49; DB 21; Length 7218;					
Best Local Similarity	2.0%; Pred. No. 1.72e-13;					
Matches	6; Conservative 165; Mismatches 122; Indels 0; Gaps 0;					

```
repeat_region /rpt_family="Alu"
3707..3728
/notes="(A)22"
/rpt_type-tandem
/rpt_unit-A
4366..4661
/rpt_family="Alu"
5327..5602
/rpt_family="Alu"
6586..6956
/rpt_family="L1"
6647..6684
/notes="(CA)19"
/rpt_type-tandem
/rpt_unit-CA
7113..7373
/rpt_family="Alu"
complement(7830..8185)
/standard_name="possible repeat"
8258..8503
/rpt_family="Alu"
9070..9387
/rpt_family="Alu"
complement(9740..9845)
/rpt_family="MER42"
complement(10440..11015)
/rpt_family="Alu"
11950..12250
/rpt_family="Alu"
12057..12085
/notes="(A)29"
/rpt_type-tandem
/rpt_unit-A
12365..12645
/rpt_family="Alu"
13727..13750
/notes="(AC)12"
/rpt_type-tandem
/rpt_unit-AC
13783..14024
/rpt_family="L1"
14175..14470
/rpt_family="Alu"
complement(14906..15259)
/standard_name="possible repeat"
15300..15613
/rpt_family="Alu"
16671..16690
/notes="(T)20"
/rpt_type-tandem
/rpt_unit-T
complement(16675..16977)
/rpt_family="Alu"
complement(16993..17085)
/rpt_family="MER42"
complement(17678..18276)
/rpt_family="Alu"
19305..19583
/rpt_family="Alu"
19914..19945
/notes="(GTTT)8"
/rpt_type-tandem
/rpt_unit-GTTT
complement(19943..20222)
/rpt_family="Alu"
21202..21496
/rpt_family="Alu"
21736..22035
/rpt_family="Alu"
22017..22038
/notes="(A)22"
/rpt_type-tandem
/rpt_unit-A

repeat_region /rpt_family="Alu"
22321..22457
complement(22638..22981)
/standard_name="possible repeat"
23022..23326
/rpt_family="Alu"
23473..23761
/rpt_family="Alu"
23744..23767
/notes="(A)24"
/rpt_type-tandem
/rpt_unit-A
complement(24639..24694)
/rpt_family="MER42"
complement(25349..25713)
/rpt_family="Alu"
complement(25727..26471)
/rpt_family="Alu"
27191..27477
/rpt_family="Alu"
27774..28057
/rpt_family="Alu"
28040..28066
/notes="(A)27"
/rpt_type-tandem
/rpt_unit-A
28712..28930
/rpt_family="MER20"
complement(28769..28838)
/notes="GRAIL 2 excellent exon, frame 0"
28987..29214
/rpt_family="Alu"
join(29485..29559,29598..29779)
/notes="95% identity dbSTS:G14522 (SHGC-11312)"
29495..29976
/standard_name="AI027942"
/notes="100% identity EST ov84a10.x1"
complement(30401..30536)
/notes="GRAIL 2 excellent exon, frame 2"
complement(30682..30733)
/rpt_family="MIR"
complement(31573..31724)
/notes="GRAIL 2 excellent exon, frame 1"
complement(32159..32232)
/notes="GRAIL 2 excellent exon, frame 2"
32388..32488
/rpt_family="MLT1"
32617..32908
/rpt_family="Alu"
32977..33088
/rpt_family="MLT1"
complement(33670..33785)
/rpt_family="Alu"
complement(34021..34144)
/notes="GRAIL 2 excellent exon, frame 2"
complement(35238..35331)
/notes="GRAIL 2 excellent exon, frame 1"
36392..36663
/rpt_family="Alu"
36901..37222
/notes="100% identity EST ou55c09.x1"
/db_xref="dbEST:AI025011"
36901..37164
/standard_name="STS-G26554"
/db_xref="dbSTS:G26554"
complement(37595..37654)
/notes="GRAIL 2 excellent exon, frame 0"
complement(join(38063..38218,38462..38578,38741..38995,
39071..39205,39532..39630,39935..40048,40300..40410,
40503..40661,41868..41972,42103..42225,42492..42569,
44379..44507))
/standard_name="histidyl-tRNA synthetase"
/notes="78%-100% protein identity GenPept:U18937"
```



```
/standard_name="RLF"
/note="65% & 69% protein identity GenPept:U22377"
3431. 3724
/rpt_family="Alu"
repeat_region
3707. 3728
/note="(A)22"
/rpt_type-tandem
/rpt_unit-A
repeat_region
4366. 4661
/rpt_family="Alu"
repeat_region
5327. 5602
/rpt_family="Alu"
repeat_region
6586. 6956
/rpt_family="L1"
repeat_region
6647. 6684
/note="(CA)19"
/rpt_type-tandem
/rpt_unit-CA
7113. 7373
/rpt_family="Alu"
repeat_region
complement(7830. 8185)
/standard_name="possible repeat"
8258. 8503
/rpt_family="Alu"
repeat_region
9070. 9387
/rpt_family="Alu"
repeat_region
complement(9740. 9845)
/rpt_family="MER42"
repeat_region
complement(10440. 11015)
/rpt_family="Alu"
repeat_region
11950. 12250
/rpt_family="Alu"
repeat_region
12057. 12085
/note="(A)29"
/rpt_type-tandem
/rpt_unit-A
12365. 12645
/rpt_family="Alu"
repeat_region
13727. 13750
/note="(AC)12"
/rpt_type-tandem
/rpt_unit-AC
13783. 14024
/rpt_family="L1"
repeat_region
14175. 14470
/rpt_family="Alu"
repeat_region
complement(14906. 15259)
/standard_name="possible repeat"
15300. 15613
/rpt_family="Alu"
repeat_region
16671. 16690
/note="(T)20"
/rpt_type-tandem
/rpt_unit-T
repeat_region
complement(16675. 16977)
/rpt_family="Alu"
repeat_region
complement(16993. 17085)
/rpt_family="MER42"
repeat_region
complement(17678. 18276)
/rpt_family="Alu"
repeat_region
19305. 19583
/rpt_family="Alu"
repeat_region
19914. 19945
/note="(GTTT)8"
/rpt_type-tandem
/rpt_unit-GTTT
repeat_region
complement(19943. 20222)
/rpt_family="Alu"
repeat_region
21202. 21496
/rpt_family="Alu"
repeat_region
21736. 22035
/rpt_family="Alu"
repeat_region
22017. 22038
```

```
/note="(A)22"
/rpt_type-tandem
22321. 22457
/rpt_family="Alu"
repeat_region
complement(22638. 22981)
/standard_name="possible repeat"
23022. 23326
/rpt_family="Alu"
repeat_region
23473. 23761
/rpt_family="Alu"
repeat_region
23744. 23767
/note="(A)24"
/rpt_type-tandem
/rpt_unit-A
complement(24639. 24694)
/rpt_family="MER42"
repeat_region
complement(25349. 25713)
/rpt_family="Alu"
repeat_region
complement(25727. 26471)
/rpt_family="Alu"
repeat_region
27191. 27477
/rpt_family="Alu"
repeat_region
27774. 28057
/rpt_family="Alu"
repeat_region
28040. 28066
/note="(A)27"
/rpt_type-tandem
/rpt_unit-A
28712. 28930
/rpt_family="MER20"
repeat_region
complement(28769. 28838)
/note="GRAIL 2 excellent exon, frame 0"
28987. 29214
/rpt_family="Alu"
misc_feature
join(29485. 29559,29598. 29779)
/note="95% identity dbSTS:G14522 (SHGC-11312)"
29495. 29976
/standard_name="AI027942"
/note="100% identity EST ov84a10.x1"
complement(30401. 30536)
/note="GRAIL 2 excellent exon, frame 2"
complement(30682. 30733)
/rpt_family="Mir"
repeat_region
complement(31573. 31724)
/note="GRAIL 2 excellent exon, frame 1"
complement(32159. 32232)
/note="GRAIL 2 excellent exon, frame 2"
32388. 32488
/rpt_family="MLT1"
repeat_region
32617. 32908
/rpt_family="Alu"
repeat_region
32977. 33088
/rpt_family="MLT1"
repeat_region
complement(33670. 33785)
/rpt_family="Alu"
misc_feature
complement(34021. 34144)
/note="GRAIL 2 excellent exon, frame 2"
complement(35238. 35331)
/note="GRAIL 2 excellent exon, frame 1"
36392. 36663
/rpt_family="Alu"
repeat_region
36901. 37222
/note="100% identity EST ou55c09.x1"
/db_xref="dbEST:AI025011"
36901. 37164
/standard_name="STSG-9983"
/db_xref="dbSTS:G28554"
complement(37595. 37654)
/note="GRAIL 2 excellent exon, frame 0"
complement(join(38063. 38218,38462. 38578,38741. 38995,
39071. 39205,39532. 39630,39935. 40048,40300. 40410,
40503. 40661,41868. 41972,42103. 42225,42492. 42569,
```

44379. .44507))  
 /standard\_name="histidyl-trna synthetase"  
 /note="78%-100% protein identity GenPept:U18937"  
 complement(38069. .38215)  
 /note="GRAIL 2 excellent exon, frame 0"  
 complement(38462. .38578)  
 ....  
 Note: remainder of annotations omitted.

Query Match 1.2%; Score 34; DB 26; Length 74371;  
 Best Local Similarity 15.5%; Pred. No. 2.75e-04;  
 Matches 13; Conservative 46; Mismatches 25; Indels 0; Gaps 0;

Db 15918 KCKFTSTKSCMSRSGKRGVSWKYYKRCAMWMTCKSKCWCWSYRBMKCYSCSYC 15977

Cp 2609 TCGCGGTAGCAGCGCTCAGACGAGCTGGAGGATGGAGCGCTCGGGATCCCTGG 2550

Db 15978 YCSSGKKYWCRCMSMYWYTCYSYK 16001

Cp 2549 CACTGGACGTAGACTTCCCTG 2526

RESULT 13  
 LOCUS A10158 201 bp DNA PAT 01-DEC-1993  
 DEFINITION S.-griseus gene for preprotease (partial).  
 ACCESSION A10158  
 NID 9490196

KEYWORDS Streptomyces griseus.  
 ORGANISM Streptomyces griseus  
 Eubacteria; Firmicutes; Actinomycetes; Streptomycetes;  
 Streptomycetaceae; Streptomycetes.  
 REFERENCE 1 (bases 1 to 201)  
 AUTHORS Garvin,R.T. and James,E.  
 TITLE Production of active proteins containing cystine residues  
 JOURNAL Patent: EP 0222279-A 2 20-MAY-1987;  
 Cangene Corporation

FEATURES  
 source  
 1. .201  
 /organism="Streptomyces griseus"  
 /db\_xref="taxon:1911"  
 1. .>201  
 /codon\_start=1  
 /transl\_table=11  
 /product="preprotease"  
 /db\_xref="PID:g490197"  
 /translation="MPHSPVSPAEPAPQPRPVVSRRLLEGGAAVLGALALSASP  
 LTAQAVRRRAADEPPEWDFGV"  
 BASE COUNT 13 a 66 c 54 g 22 t 46 others  
 ORIGIN

Query Match 1.1%; Score 30; DB 21; Length 201;  
 Best Local Similarity 36.1%; Pred. No. 4.44e-02;  
 Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

Db 2 TGCSCACTCCSCSTSCSCGSGAGTSCGSCGCCAGCCGCGGCGCCGCGC 61

Qy 26 TCGTCCGACGTTGGAGAGCCCTGGCCGCCACCCCGCGATGCCGCTCCCGCT 85

Db 62 CSGTSGTSCGCGCGCTCTSGAGGSGGSGCGSGTSGTSGGCGCTSGCSTST 121

Qy 86 GCCGAGCGTGGCTCCCTGCTGGCAGCCACTACCGGAGGTGCTGCCGCTGCCACGT 145

Db 122 CSGCSTCCSCSTSCAGGCCAGG 145

Qy 146 TCGTCCGCGCTGGGGCCCCCAGG 169

RESULT 14  
 LOCUS A10161 201 bp DNA PAT 01-DEC-1993  
 DEFINITION Synthetic DNA for preprotease leader & prochymosin.  
 ACCESSION A10161  
 NID 9490555

KEYWORDS unidentified.  
 SOURCE unclassified.  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 201)  
 AUTHORS Garvin,R.T. and James,E.  
 TITLE Production of active proteins containing cystine residues  
 JOURNAL Patent: EP 0222279-A 5 20-MAY-1987;  
 Cangene Corporation

FEATURES  
 source  
 1. .201  
 Location/Qualifiers

CDS  
 1. .>201  
 /organism="Artificial sequences"  
 /codon\_start=1  
 /product="preprotease leader & prochymosin"  
 /db\_xref="PID:g490656"  
 /translation="MPHSPVSPAEPAPQPRPVVSRRLLEGGAAVLGALALSASP  
 LTAQAVRRRAAREITRPLYKGR"  
 BASE COUNT 16 a 66 c 51 g 22 t 46 others  
 ORIGIN

Query Match 1.1%; Score 30; DB 21; Length 201;  
 Best Local Similarity 36.1%; Pred. No. 4.44e-02;  
 Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

Db 2 TGCSCACTCCSCSTSCSCGSGAGTSCGSCGCCAGCCGCGGCGCCGCGC 61

Qy 26 TCGTCCGACGTTGGAGAGCCCTGGCCGCCACCCCGCGATGCCGCTCCCGCT 85

Db 62 CSGTSGTSCGCGCGCTCTSGAGGSGGSGCGSGTSGTSGGCGCTSGCSTST 121

Qy 86 GCCGAGCGTGGCTCCCTGCTGCCAGCCACTACCGGAGGTGCTGCCGCTGCCACGT 145

Db 122 CSGCSTCCSCSTSCAGGCCAGG 145

Qy 146 TCGTCCGCGCTGGGGCCCCCAGG 169

RESULT 15  
 LOCUS A10162 201 bp DNA PAT 01-DEC-1993  
 DEFINITION Synthetic DNA for preprotease leader & prochymosin.  
 ACCESSION A10162  
 NID 9489115  
 KEYWORDS synthetic construct.  
 ORGANISM synthetic construct  
 REFERENCE 1 (bases 1 to 201)  
 AUTHORS Garvin,R.T. and James,E.  
 TITLE Production of active proteins containing cystine residues  
 JOURNAL Patent: EP 0222279-A 6 20-MAY-1987;  
 Cangene Corporation

FEATURES  
 source  
 1. .201  
 Location/Qualifiers

CDS  
 1. .>201  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 BASE COUNT 22 a 51 c 66 g 16 t 46 others  
 ORIGIN

Query Match 1.1%; Score 30; DB 21; Length 201;  
 Best Local Similarity 36.1%; Pred. No. 4.44e-02;  
 Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

Db 57 CCTGGCGCTSGAGGSGGSGAGSGAGGCGCCCGCCAGCCGCGGCTAGTGGCTGC 116

Cp 169 CCTGGGCCCCAGCGCGCGCACGACGCTGGCCAGCGGAGCAGCCTCGGGTAGTGGCTGC 110

Db 117 CSAGSAGSGCGSGASACSGAGGSGGSGCGSGGCGGCGGCGGCGGCGGCGGCGGCGG 176

Cp 109 GCAGCAGGAGCGCAGCGCTCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 50

Db 177 CSGSGASACSGGSGAGTGGGCA 200

Cp 49 CCAGGGCTTCCACGCTGCGCAGCA 26

Search completed: Fri Dec 25 12:03:06 1998  
Job time : 11319 secs.

\*\*\*\*\*

W P E R L F (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 12:03:25 1998; MasPar time 8218.69 Seconds  
Tabular output not generated. 620.553 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1-2848) from US08951733.seq (1 of 3)  
Perfect Score: 2848  
N.A. Sequence: 1 CAGCGGTCGCGGAGCGCTG.....GATCGCGGCGCCAGCGCTAT 2848  
Comp: GTGCGGAGCGCGCTGCGGAG.....CTACGCGCGGTCGCGGATA

Scoring table: TABLE default  
Gap open 30; Gap extend 1  
Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est55  
Database: genbank-est107  
5:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13  
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17  
14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21  
18:gb-est21 19:gb-est23 20:gb-est4 21:gb-est5 22:gb-est6  
23:gb-est7 24:gb-est8 25:gb-est9 26:gb-est1 27:gb-est2  
28:gb-est3 29:gb-est4

Statistics: Mean 11.733; Variance 2.670; scale 4.395

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	385	13.5	389	8	AA281296	EST08902.r1 NCI_CGAP_GC
2	187	6.6	409	24	AA311750	EST182469 Jurkat T-cel
3	58	2.0	252	12	AA754459	97SN1787 Rice Immature
4	54	1.9	252	12	AA754459	97SN1787 Rice Immature
5	49	1.7	247	12	AA754458	97SN1784 Rice Immature
6	44	1.5	247	12	AA754458	97SN1784 Rice Immature
7	33	1.2	375	23	AA200728	mul3h09.r1 Soares 2BDM
8	33	1.2	2275	11	AF034173	Homo sapiens ntcon6 co
9	31	1.1	660	11	AF034177	Homo sapiens ntcon6 co
10	28	1.0	1287	12	AF038250	Homo sapiens clone ntc
11	28	1.0	1287	12	AF038250	Homo sapiens clone ntc
12	25	0.9	143	18	AA171394	EST217350 Normalized r

13	25	0.9	181	6	AA386387	EST81369 Prostate glan
14	25	0.9	187	13	AA855630	Y068h11.r1 Stratagene
15	25	0.9	213	16	R72797	Y109c09.r1 Homo sapien
16	25	0.9	242	15	AA1012107	EST206558 Normalized r
17	25	0.9	278	13	AA882818	TENS0393 T. cruzi epim
18	25	0.9	301	14	AA848961	EST191723 Normalized r
19	25	0.9	302	15	AA1010844	EST205295 Normalized r
20	27	0.9	317	27	AQ080007	CIT-HSP-2288C3.TF CIT-
21	25	0.9	318	14	AA799774	EST189271 Normalized r
22	25	0.9	318	16	H50134	Y027a07.r1 Homo sapien
23	25	0.9	336	18	AA1172396	EST218402 Normalized r
24	26	0.9	343	22	AA098755	T3985 MVAT4 bloodstrea
25	25	0.9	359	24	AA303595	EST16305 Aorta endothe
26	25	0.9	382	5	RS4656	YJ74d04.r1 Homo sapien
27	26	0.9	396	19	T26788	T529 Trypanosoma bruce
28	26	0.9	400	16	H44707	YP24c08.r1 Homo sapien
29	25	0.9	420	17	AT071447	UI-R-C1-ku-c-05-0-UI.s
30	25	0.9	438	14	AA924206	UI-R-Al-dx-a-11-0-UI.s
31	25	0.9	478	14	AA925372	UI-R-Al-ee-d-07-0-UI.s
32	25	0.9	487	16	R86860	Ym86a03.r1 Homo sapien
33	25	0.9	504	15	AT101020	EST204571 Normalized r
34	25	0.9	511	17	AT072026	UI-R-C2-nd-e-12-0-UI.s
35	25	0.9	511	15	AA997735	UI-R-C0-hn-b-08-0-UI.s
36	25	0.9	540	14	AA901048	UI-R-Al-dv-e-09-0-UI.s
37	25	0.9	590	10	AA622227	no43c05.sl NCI_CGAP.Pr
38	27	0.9	660	11	AF034177	Homo sapiens ntcon6 co
39	25	0.9	677	18	AA105428	EST214717 Normalized r
40	25	0.9	705	14	AA941159	LD5143.5Prime LD Dros
41	26	0.9	7275	11	AF034173	Homo sapiens ntcon2 co
42	24	0.8	299	21	W04101	T1794 MVAT4 bloodstrea
43	24	0.8	398	5	T47934	Yb18d04.sl Homo sapien
44	24	0.8	456	6	CI9919	Rice cDNA, partial seq
45	24	0.8	901	20	W21990	60h10 Human retina CDN

ALIGNMENTS

RESULT 1  
LOCUS AA281296 389 bp mRNA EST 14-AUG-1997  
DEFINITION zt08902.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:712562 5',  
mRNA sequence.  
ACCESSION AA281296  
NID G1924194  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;  
Homo.  
REFERENCE 1 (bases 1 to 389)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITILE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 2187 Std Error: 0.00  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 385.  
Location/Qualifiers  
1. .389  
/organism="Homo sapiens"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD+),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTT-3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

/db\_xref="taxon:9606"  
/clone="IMAGE:712562"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
<1. >389

BASE COUNT 87 a 102 c 123 g 77 t

ORIGIN

Query Match 13.5%; Score 385; DB 8; Length 389;  
Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGTCTTTCTTT 60  
QY 1690 GCCAAGTTCCTGCTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGTCTTTCTTT 1749

Db 61 TATGTCACGGAGACACAGCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGG 120  
QY 1750 TATGTCACGGAGACACAGCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGG 1809

Db 121 AGCAAGTTGCAAGCATTGGAATCAGACAGCACATTGAAGAGGGTGCAGCTCGGGAGCTG 180  
QY 1810 AGCAAGTTGCAAGCATTGGAATCAGACAGCACATTGAAGAGGGTGCAGCTCGGGAGCTG 1869

Db 181 TCGGAGCAGAGGTGAGCAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTC 240  
QY 1870 TCGGAGCAGAGGTGAGCAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTC 1929

Db 241 CGCTTCATCCCAAGCCTGAGCGGCTGCGCGGATGTGACATGGACTAGCTCGTGGGA 300  
QY 1930 CGCTTCATCCCAAGCCTGAGCGGCTGCGCGGATGTGACATGGACTAGCTCGTGGGA 1989

Db 301 GCCAAGCTTCCCGAGAAAGAGGCGCGAGCGCTCTCACCTCGAGGGTGAAGGCACGTG 360  
QY 1990 GCCAAGCTTCCCGAGAAAGAGGCGCGAGCGCTCTCACCTCGAGGGTGAAGGCACGTG 2049

Db 361 TTCAGCTGCTCACTAGCAGCGGCGCG 389  
QY 2050 TTCAGCTGCTCACTAGCAGCGGCGCG 2078

RESULT 2 AA311750 409 bp mRNA EST 19-APR-1997  
LOCUS EST182469 Jurkat T-cells Vi Homo sapiens cDNA 5' end, mRNA  
DEFINITION sequence.  
ACCESSION AA311750  
NID g1964077  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 409)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl), 3-174 (1995)  
96026280

## JOURNAL

## MEDLINE

## COMMENT

Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/ngi/ngi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. .409  
/organism="Homo sapiens"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"  
/db\_xref="ATCC (inhost):158964"  
/db\_xref="taxon:9606"  
/clone\_lib="Jurkat T-cells VI"  
/cell\_type="T-lymphocyte"  
<1. >409

## FEATURES

## source

BASE COUNT 65 a 120 c 133 g 86 t 5 others  
ORIGIN

Query Match 6.6%; Score 187; DB 24; Length 409;  
Best Local Similarity 98.9%; Pred. No. 5.02e-269;  
Matches 187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GTTTGGTGGATGATTTCTTTGTTGGTGCACCTCACCTCACCCAGCGAAAACCTTCCTCA 60  
QY 2660 GTTTGGTGGATGATTTCTTTGTTGGTGCACCTCACCTCACCCAGCGAAAACCTTCCTCA 2719

Db 61 GGACCTGGTCCGAGGTGCTCCCTGAGTATGGCTCGGTGGTGAACCTTCGGAGACAGTGG 120  
QY 2720 GGACCTGGTCCGAGGTGCTCCCTGAGTATGGCTCGGTGGTGAACCTTCGGAGACAGTGG 2779

Db 121 TGAACCTCCCTGTAGAAGACGAGCCCTGGGTGGCAGCGCTTTTTCAGATCGCGGCC 180  
QY 2780 TGAACCTCCCTGTAGAAGACGAGCCCTGGGTGGCAGCGCTTTTTCAGATCGCGGCC 2839

Db 181 ACGCCTAT 189  
QY 2840 ACGCCTAT 2848

RESULT 3 AA754459 252 bp mRNA EST 20-JAN-1998  
LOCUS 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa  
DEFINITION cDNA clone 97SN1787, mRNA sequence.  
ACCESSION AA754459  
NID g2801165  
KEYWORDS EST.  
SOURCE rice.  
ORGANISM Oryza sativa

REFERENCE 1 (bases 1 to 252)  
AUTHORS Nahn,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P., Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y., Lee,M.C. and Eun,M.Y.  
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed Unpublished (1998)

REFERENCE 1 (bases 1 to 252)  
AUTHORS Nahn,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P., Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y., Lee,M.C. and Eun,M.Y.  
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed Unpublished (1998)







```
Query Match 1.2%; Score 33; DB 23; Length 375;
Best Local Similarity 75.4%; Pred. No. 5.05e-13;
Matches 49; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 144 AGAGCATCTATGAATGAGAGCAGCAGCCTGTTTGACTTCTTCTGCACTTCCTGC 203
||||| ||| ||||||| ||||||| ||||| ||| ||||||| ||||||| ||||| ||
QY 2447 AGAGCTCTCTGATGAGGCCAGCAGTGGCTTTCGACGCTCTCTACGCTTCATGT 2506
Db 204 GTCAC 208
QY 2507 GCCAC 2511

RESULT 8 AF034173 2275 bp mRNA EST 22-DEC-1997
LOCUS Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
DEFINITION AF034173
ACCESSION AF034173
NID 92707735
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Generation of a transcription map in the region immediately
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2275)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 2275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon2 contig"
/tissue_type="liver; brain"
/dev_stage="fetus"
/note="similar to Br140"

BASE COUNT 438 a 619 c 470 g 599 t 149 others
ORIGIN

Query Match 1.2%; Score 33; DB 11; Length 2275;
Best Local Similarity 13.3%; Pred. No. 5.05e-13;
Matches 11; Conservative 47; Mismatches 25; Indels 0; Gaps 0;

Db 1523 GKKKCKYKRYKKTSTYKSWRWYTYTYWCCTSMKASGASCMRWGMYGSR 1582
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2648 GGCTGCTCTGGTGTGGTGGATGATTTCTTGTGGTGCACCTCACCACCGGGA 2707
Db 1583 SRSRWYWGWSMCGYMTKRY 1605
QY 2708 AAACCTTCTCAGGACCTCGTC 2730

RESULT 9 AF034177 660 bp mRNA EST 22-DEC-1997
LOCUS Homo sapiens ntcon6 contig mRNA, partial sequence, mRNA sequence.
DEFINITION AF034177
ACCESSION AF034177
NID 92707739
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Generation of a transcription map in the region immediately
```

```
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 660)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon6 contig"
/tissue_type="fetal brain; fetal liver; adult muscle"
/note="similar to CutA"

BASE COUNT 162 a 161 c 173 g 109 t 55 others
ORIGIN

Query Match 1.1%; Score 31; DB 11; Length 660;
Best Local Similarity 15.6%; Pred. No. 1.45e-10;
Matches 7; Conservative 31; Mismatches 7; Indels 0; Gaps 0;

Db 384 WSSRCYRSYGYHYCYKCDMSCTTSKSRWYKSRWCD 428
:::| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 AGAAGTGTGGCTTGGCTTGGCTGCTGCGGCGGCGG 391

RESULT 10 AF038250 1287 bp mRNA EST 29-JAN-1998
LOCUS Homo sapiens clone ntcon9 mRNA, partial sequence, mRNA sequence.
DEFINITION AF038250
ACCESSION AF038250
NID 92815880
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1287)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
source
1. 1287
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="ntcon9"
/dev_stage="fetal"
/tissue_type="brain; liver"
/note="similar to HSRP20"

BASE COUNT 349 a 219 c 293 g 361 t 65 others
ORIGIN

Query Match 1.0%; Score 28; DB 12; Length 1287;
Best Local Similarity 29.1%; Pred. No. 4.79e-07;
Matches 23; Conservative 29; Mismatches 27; Indels 0; Gaps 0;

Db 401 DCBTGGDTBHBVCBNCBANADGATBACGKGVGBGASTCMTHYNCDCCKTCGSAG 460
:|:|:| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 GCCTGGTGTGCTGCGCTGGGAGCAGCGCCGCCGCCCTCTCTCCGCCAGG 286
Db 461 TVNHHDWSMAGARAGCT 479
QY 287 TGTCTGCTCGAAGGAGCT 305

RESULT 11 AF038250 1287 bp mRNA EST 29-JAN-1998
LOCUS Homo sapiens clone ntcon9 mRNA, partial sequence, mRNA sequence.
DEFINITION
```

```

ACCESSION AF038250
NID 92815880
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1287)
AUTHORS Tripodis, N. and Ragoussis, J.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES
    source
        1..1287
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="6"
            /map="6p21.3"
            /clone="ntcon9"
            /dev_stage="fetal"
            /tissue_type="brain; liver"
            /note="similar to HSRP20"
BASE COUNT 349 a 219 c 293 g 361 t 65 others
ORIGIN
Query Match 1.0%; Score 28; DB 12; Length 1287;
Best Local Similarity 23.8%; Pred. No. 4.79e-07;
Matches 15; Conservative 29; Mismatches 19; Indels 0; Gaps 0;
Db 407 GDTBSHVCMCBANADGATBATCGKGVGASCTMTHYNCCDCKTCGSAGVTVNHHD 466
Cp 1426 GCGGAGCAGCTCACCGACGCGGGGTCGTGTCCTCCCTCGGGGCGGCCACAG 1367
Db 467 WSM 469
Cp 1366 AGC 1364

RESULT 12
LOCUS A1171394 143 bp mRNA EST 07-OCT-1998
DEFINITION EST217350 Normalized rat muscle, Bento Soares Rattus sp. cDNA clone
RMUBK19 3' end, mRNA sequence.
ACCESSION A1171394
NID 93711434
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 143)
AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
TITLE Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Other_ESTs: TC52270
Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers
    1..143
        /organism="Rattus sp."
        /note="Organ: muscle; Vector: p7T3Pac; Site_1: EcoRI;
        Site_2: NotI"
        /db_xref="taxon:10118"
        /clone_lib="RMUBK19"
        /clone_lib="Normalized rat muscle, Bento Soares"
BASE COUNT 14 a 49 c 58 g 22 t

```

```

ORIGIN
Query Match 0.9%; Score 25; DB 18; Length 143;
Best Local Similarity 76.6%; Pred. No. 8.99e-04;
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Db 4 GGGAGCGATGGCGGCTGCGGGGCTCGCGCGCGCTCCCGCTGGTG 50
Cp 103 GGGAGCGCGGCTGCGGAGCGGAGCGCGCGCATCGCGGGGTG 57

RESULT 13
LOCUS AA386387 181 bp mRNA EST 22-APR-1997
DEFINITION EST81369 Prostate gland 1 Homo sapiens cDNA 5' end similar to
prolyl 4-hydroxylase, beta subunit/protein disulfide
isomerase/thyroid hormone-binding protein, mRNA sequence.
ACCESSION AA386387
NID G2038747
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;
Homo.
REFERENCE 1 (bases 1 to 181)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
Fraser, C.M. and Venter, J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Other_ESTs: EST81368 THC170792
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
    1..181
        /organism="Homo sapiens"
        /note="Organ: prostate; Vector: pBluescript SK-; Site_1:
        EcoRI; Site_2: XhoI"
        /db_xref="ATCC (inhost):174439"
        /db_xref="taxon:9606"
        /clone_lib="Prostate gland 1"
        /sex="male"
        /dev_stage="adult, 21 yrs"
        /cl. >181
BASE COUNT 25 a 60 c 59 g 32 t 5 others
ORIGIN
Query Match 0.9%; Score 25; DB 6; Length 181;

```

```

ACCESSION R72797
UNID 9846829
KEYWORDS EST.
SOURCE human clone=157744 library=Soares breast 2NDHBst vector=pt7T3D
        (Pharmacia) with a modified polylinker host=DH10B (ampicillin
        resistant) primer=M3RP1 Rsite1-Not I Rsite2-Eco RI Adult female.
        1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-
        TGTACCAATCTGAAGTGGAGGCGCCGCTTTTTTTTTTTT 3'],
        double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
        digested with Not I and cloned into the Not I and Eco RI sites of a
        modified pT7m3 vector (Pharmacia). Library went through one round
        of normalization to a Cot = 230. Library constructed by Bento
        Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens
        Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata;
        Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
        Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
        Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
        1 (bases 1 to 213)
        Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
        Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
        Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
        Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
        Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 185
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
        source
            1..213
                /organism="Homo sapiens"
                /clone="157744"

BASE COUNT 40 a 72 c 69 g 32 t
ORIGIN

Query Match 0.9%; Score 25; DB 16; Length 213;
Best Local Similarity 74.5%; Pred. No. 8,99e-04;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Ddb 32 CTGCCCGGCGCACGTTCTGTGACGCGTGAGACCGGAGAGCTCTCGGCTG 82
QY 130 CTGCCCTGGCCACGTCGTGTGGCGGCCCTGTGGGCCCCCAGGCTGGCGGCTG 180

Search completed: Fri Dec 25 14:50:06 1998
Job time : 10001 secs.

```

Search completed: Fri Dec 25 14:50:06 1998  
Job time : 10001 secs.

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**







[illegible]



[illegible]

DE Generic DNA sequence to generate a random TSAR peptide library.  
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
 KW effector domain; concatenated heterofunctional protein; linker;  
 KW direct; rapid; detection; screening; treatment; generic; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 55..60  
 FT /tag= a  
 FT /note= "this sequence represents 'Z'; Z can be a  
 FT sequence of 6, 9 or 12 nucleotides (see  
 FT comments)"  
 PN WO9418318-A.  
 PD 18-AUG-1994.  
 PE 01-FEB-1994; U00977.  
 PR 01-FEB-1993; US-013416.  
 PR 30-DEC-1993; US-176500.  
 PR 31-JAN-1994; US-189331.  
 PA (UUNC-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Kay BK;  
 DR WPI: 94-279739/34.  
 DR P-PSDB; R65150 and R65151.  
 DR Identifying proteins or peptide(s) which bind a ligand - by  
 PT screening a recombinant vector library expressing fusion proteins  
 PT comprising a binding domain and an effector domain  
 PS Disclosure; Page 35; 255pp; English.  
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally  
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
 CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X  
 CC and Y are flanking restriction sites (X is not the same as Y) that are  
 CC not specified further. Other generic sequences are shown in Q70466-68.  
 CC Other specific peptides generated by these generic sequences are shown in  
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
 CC comprising at least two functional regions - a binding domain with  
 CC affinity for a ligand and a second effector peptide portion that is  
 CC chemically or biologically active. They may further comprise a linker  
 CC peptide between the 2 domains. The oligonucleotides are also designed so  
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
 CC in, or flanking, the unpredicted or variant residues. These residues  
 CC confer some degree of conformational rigidity to the peptides. The TSARs  
 CC or compns. comprising a TSAR binding domain can be used in vivo to  
 CC deliver a chemically or biologically active moiety, eg. metal ion,  
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
 CC cell. They can also replace the function of macromolecules, eg.  
 CC monoclonal or polyclonal antibodies and therefore circumvent the need  
 CC for complex methods of hybridoma formation or in vivo antibody  
 CC production. The TSARs are easily characterised and have designed  
 CC activity allowing direct and rapid detection in a screening process.  
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
 Query Match 1.3%; Score 38; DB 12; Length 114;  
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;  
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;  
 Db 3 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 52  
 Cp 1527 GAAGCGGCTGCTGTGCTGAGAGCCAGAGCGCTGGGGACACCGCGGCGAGCA 1468  
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114  
 Cp 1467 GCCCGCAGCAGCGGTACACCTGCAGCGGCTGCTGTGCTGGCGGACAGC 1416  
 RESULT 13  
 ID Q70468 standard; DNA; 114 BP.  
 AC Q70468;  
 DE 05-APR-1995 (first entry)  
 DE Generic DNA sequence to generate a random TSAR peptide library.  
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
 KW effector domain; concatenated heterofunctional protein; linker;  
 KW direct; rapid; detection; screening; treatment; generic; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 55..60  
 FT /tag= a  
 FT /note= "this sequence represents 'Z'; Z can be a  
 FT sequence of 6, 9 or 12 nucleotides (see  
 FT comments)"  
 PN WO9418318-A.  
 PD 18-AUG-1994.  
 PE 01-FEB-1994; U00977.  
 PR 01-FEB-1993; US-013416.  
 PR 30-DEC-1993; US-176500.  
 PR 31-JAN-1994; US-189331.  
 PA (UUNC-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Kay BK;  
 DR WPI: 94-279739/34.  
 DR P-PSDB; R65150 and R65151.  
 DR Identifying proteins or peptide(s) which bind a ligand - by  
 PT screening a recombinant vector library expressing fusion proteins  
 PT comprising a binding domain and an effector domain  
 PS Disclosure; Page 35; 255pp; English.  
 CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally  
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
 CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X  
 CC and Y are flanking restriction sites (X is not the same as Y) that are  
 CC not specified further. Other generic sequences are shown in Q70466-68.  
 CC Other specific peptides generated by these generic sequences are shown in  
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
 CC comprising at least two functional regions - a binding domain with  
 CC affinity for a ligand and a second effector peptide portion that is  
 CC chemically or biologically active. They may further comprise a linker  
 CC peptide between the 2 domains. The oligonucleotides are also designed so  
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
 CC in, or flanking, the unpredicted or variant residues. These residues  
 CC confer some degree of conformational rigidity to the peptides. The TSARs  
 CC or compns. comprising a TSAR binding domain can be used in vivo to  
 CC deliver a chemically or biologically active moiety, eg. metal ion,  
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
 CC cell. They can also replace the function of macromolecules, eg.  
 CC monoclonal or polyclonal antibodies and therefore circumvent the need  
 CC for complex methods of hybridoma formation or in vivo antibody  
 CC production. The TSARs are easily characterised and have designed  
 CC activity allowing direct and rapid detection in a screening process.  
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
 Query Match 1.3%; Score 38; DB 12; Length 114;  
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;  
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;  
 Db 3 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 52  
 Cp 1527 GAAGCGGCTGCTGTGCTGAGAGCCAGAGCGCTGGGGACACCGCGGCGAGCA 1468  
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114  
 Cp 1467 GCCCGCAGCAGCGGTACACCTGCAGCGGCTGCTGTGCTGGCGGACAGC 1416

FT /tag= a  
 FT /note= "this sequence represents 'Z'; Z can be a  
 FT sequence of 6, 9 or 12 nucleotides (see  
 FT comments)"  
 PN WO9418318-A.  
 PD 18-AUG-1994.  
 PE 01-FEB-1994; U00977.  
 PR 01-FEB-1993; US-013416.  
 PR 30-DEC-1993; US-176500.  
 PR 31-JAN-1994; US-189331.  
 PA (UUNC-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Kay BK;  
 DR WPI: 94-279739/34.  
 DR P-PSDB; R65151.  
 DR Identifying proteins or peptide(s) which bind a ligand - by  
 PT screening a recombinant vector library expressing fusion proteins  
 PT comprising a binding domain and an effector domain  
 PS Disclosure; Page 35; 255pp; English.  
 CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally  
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
 CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X  
 CC and Y are flanking restriction sites (X is not the same as Y) that are  
 CC not specified further. Other generic sequences are shown in Q70466-68.  
 CC Other specific peptides generated by these generic sequences are shown in  
 CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
 CC comprising at least two functional regions - a binding domain with  
 CC affinity for a ligand and a second effector peptide portion that is  
 CC chemically or biologically active. They may further comprise a linker  
 CC peptide between the 2 domains. The oligonucleotides are also designed so  
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
 CC in, or flanking, the unpredicted or variant residues. These residues  
 CC confer some degree of conformational rigidity to the peptides. The TSARs  
 CC or compns. comprising a TSAR binding domain can be used in vivo to  
 CC deliver a chemically or biologically active moiety, eg. metal ion,  
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
 CC cell. They can also replace the function of macromolecules, eg.  
 CC monoclonal or polyclonal antibodies and therefore circumvent the need  
 CC for complex methods of hybridoma formation or in vivo antibody  
 CC production. The TSARs are easily characterised and have designed activity  
 CC allowing direct and rapid detection in a screening process.  
 SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
 Query Match 1.3%; Score 38; DB 12; Length 114;  
 Best Local Similarity 4.5%; Pred. No. 9.79e-07;  
 Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;  
 Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62  
 Cp 1542 CTTGCTGTCTGAGAGCGGCTGCTGTGCTGAGAGCCAGAGCGCTGGGGGCAC 1483  
 Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114  
 Cp 1482 CAGCGCGGCGAGCGCGCGCGAGCGGTACACCTGCAGCGGCTGCTG 1431  
 RESULT 14  
 ID Q70467 standard; DNA; 114 BP.  
 AC Q70467;  
 DE 05-APR-1995 (first entry)  
 DE Generic DNA sequence to generate a random TSAR peptide library.  
 KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
 KW effector domain; concatenated heterofunctional protein; linker;  
 KW direct; rapid; detection; screening; treatment; generic; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 55..60  
 FT /tag= a  
 FT /note= "this sequence represents 'Z'; Z can be a  
 FT sequence of 6, 9 or 12 nucleotides (see  
 FT comments)"  
 PN WO9418318-A.  
 PD 18-AUG-1994.  
 PE 01-FEB-1994; U00977.

PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI: 94-279739/34.  
DR P-PSDB: R65153.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 255pp; English.  
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X  
CC and Y are flanking restriction sites (X is not the same as Y) that are  
CC not specified further. Other generic sequences are shown in Q70466-68.  
CC Other specific peptides generated by these generic sequences are shown in  
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
CC comprising at least two functional regions - a binding domain with  
CC affinity for a ligand and a second effector peptide portion that is  
CC chemically or biologically active. They may further comprise a linker  
CC peptide between the 2 domains. The oligonucleotides are also designed so  
CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
CC in, or flanking, the unpredicted or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARs  
CC or compsns. comprising a TSAR binding domain can be used in vivo to  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need for  
CC complex methods of hybridoma formation or in vivo antibody production.  
CC The TSARs are easily characterised and have designed activity allowing  
CC direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.3%; Score 38; DB 12; Length 114;  
Best Local Similarity 4.5%; Pred. No. 9,79e-07;  
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnnnn 62  
Cp 1557 CAGGAGATGAACCTTGGTGTCTCCTGAGGAGCGGGTTCGTGTCCTGGAGCCCCA 1498  
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 114  
Cp 1497 GAGCGCTGGGGCCACGACCGCGCGCAGCAGCGCCGCGCAGGCGGTACACC 1446

RESULT 15  
ID Q70470 standard; DNA; 114 BP.  
AC Q70470;  
DT 10-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FS Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /\*tag= a  
FT /note= "encoded by Z (see comments)"

DN W09418318-A.  
FD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI: 94-279739/34.  
DR P-PSDB: R58378.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain  
PS Disclosure; Page 36; 255pp; English.  
CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)4(CAC)(NNB)82(NNB)6(CAC)(NNB)8  
CC -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same  
CC as Y) that are not specified further. The peptides generated by this and  
CC other generic sequences (Q70471-73) have invariant histidine residues  
CC incorporated into variant sequences. TSARs are concatenated  
CC heterofunctional proteins or peptides, comprising at least two functional  
CC regions - a binding domain with affinity for a ligand and a second  
CC effector peptide portion that is chemically or biologically active. They  
CC may further comprise a linker peptide between the 2 domains. The TSARs  
CC or compsns. comprising a TSAR binding domain can be used in vivo to  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed  
CC activity allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 5 A; 10 C; 0 G; 0 T;

Query Match 1.3%; Score 36; DB 12; Length 114;  
Best Local Similarity 7.7%; Pred. No. 1.03e-05;  
Matches 8; Conservative 30; Mismatches 66; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnnnn 62  
Cp 2625 CAGCTTGTCTCCATGTCGGCGTAGCACAGGCTCAGAGCAGCGTGGAGAGATGGAGCC 2566  
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 106  
Cp 2565 CTGCGGGATCCCTGTGCACTGGAGTAGGACTTGCCTGTATGC 2522

Search completed: Fri Dec 25 15:11:06 1998  
Job time : 1239 secs.

THIS PAGE BLANK (USPTO)

\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 15:16:55 1998; MasPar time 5202.31 seconds

Tabular output not generated. 586.621 Million cell updates/sec

Title: >US-08-951-733-13

Description: (1-2848) from US08951733.seq (1 of 3)

Perfect Score: 2848

N.A. Sequence: 1 CACGCGCGCGGCGAGCGTG.....GATCGCGCGCGCGCGCTAT 2848

Comp: GTGGCAGCGCGCGTGGCGAC.....CTACGCGCGGTCGCGGATA

Scoring table: TABLE default

Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777161 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-pending  
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005  
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82  
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88  
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8  
29:NEWU9

Statistics: Mean 10.979; Variance 4.476; scale 2.453

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2848	100.0	2848	22	US-08-951-Sequence 13, Applicati	0.00e+00
2	2848	100.0	2848	22	US-08-951-Sequence 19, Applicati	0.00e+00
3	2840	99.7	7029	22	US-08-951-Sequence 1, Applicati	0.00e+00
4	2837	99.6	4015	21	US-08-854-Sequence 224, Applicat	0.00e+00
5	2837	99.6	4015	22	US-08-912-Sequence 1, Applicati	0.00e+00
6	2837	99.6	4015	24	US-09-052-Sequence 1, Applicati	0.00e+00
7	2834	99.5	4023	23	US-09-026-Sequence 35, Applicati	0.00e+00
8	2821	88.5	4029	21	US-08-854-Sequence 173, Applicat	0.00e+00
9	2521	88.5	4029	21	US-08-851-Sequence 173, Applicat	0.00e+00
10	2444	85.8	3855	22	US-08-911-Sequence 18, Applicati	0.00e+00
11	2444	85.8	3855	22	US-08-912-Sequence 4, Applicatio	0.00e+00
12	1497	52.6	4200	22	US-08-912-Sequence 6, Applicatio	0.00e+00
13	825	29.0	3496	23	US-09-042-Sequence 1, Applicati	0.00e+00
14	816	28.7	3346	23	US-09-026-Sequence 36, Applicati	0.00e+00
15	767	26.9	2176	22	US-08-912-Sequence 3, Applicatio	0.00e+00

16	735	25.8	2171	21	US-08-854-Sequence 100, Applicat	0.00e+00
17	735	25.8	2171	21	US-08-846-Sequence 100, Applicat	0.00e+00
18	735	25.8	2171	21	US-08-851-Sequence 100, Applicat	0.00e+00
19	535	18.8	535	22	US-08-911-Sequence 52, Applicati	0.00e+00
20	460	16.2	550	22	US-08-911-Sequence 54, Applicati	0.00e+00
21	385	13.5	389	21	US-08-844-Sequence 62, Applicati	0.00e+00
22	385	13.5	389	21	US-08-846-Sequence 62, Applicati	0.00e+00
23	385	13.5	389	21	US-08-851-Sequence 62, Applicati	0.00e+00
24	385	13.5	389	22	US-08-911-Sequence 17, Applicati	0.00e+00
25	385	13.5	389	21	US-08-854-Sequence 62, Applicati	0.00e+00
26	385	13.5	389	22	US-08-912-Sequence 8, Applicatio	0.00e+00
27	182	6.4	182	22	US-08-912-Sequence 9, Applicatio	7.34e-146
28	120	4.2	2651	23	US-09-042-Sequence 5, Applicatio	1.73e-84
29	97	3.4	240	22	US-08-912-Sequence 7, Applicatio	2.26e-62
30	69	2.4	477	11	US-07-904-Sequence 10, Applicati	2.34e-36
c 31	69	2.4	477	11	US-07-904-Sequence 10, Applicati	2.34e-36
32	68	2.4	578	18	US-08-569-Sequence 14, Applicati	1.87e-35
33	68	2.4	578	19	US-08-624-Sequence 14, Applicati	1.87e-35
c 34	68	2.4	578	19	US-08-624-Sequence 14, Applicati	1.87e-35
c 35	68	2.4	578	18	US-08-569-Sequence 14, Applicati	1.87e-35
36	69	2.4	1808	23	US-09-042-Sequence 4, Applicatio	2.34e-36
37	67	2.4	7218	17	US-08-466-Sequence 14, Applicati	1.48e-34
c 38	66	2.3	383	16	US-08-446-Sequence 8, Applicatio	1.17e-33
c 39	66	2.3	383	15	US-08-311-Sequence 8, Applicatio	1.17e-33
c 40	66	2.3	383	17	US-08-451-Sequence 8, Applicatio	1.17e-33
c 41	66	2.3	383	16	US-08-446-Sequence 8, Applicatio	1.17e-33
c 42	66	2.3	383	16	US-08-446-Sequence 8, Applicatio	1.17e-33
c 43	66	2.3	383	16	US-08-446-Sequence 8, Applicatio	1.17e-33
c 44	62	2.2	340	19	US-08-617-Sequence 9, Applicatio	4.27e-30
45	62	2.2	383	17	US-08-451-Sequence 8, Applicatio	4.27e-30

## ALIGNMENTS

RESULT 1  
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.  
AC xxxxxx  
DT  
Sequence 13, Application US/08951733  
Sequence 13, Application US/08951733  
GENERAL INFORMATION:  
APPLICANT: Harrington, Lea A.  
APPLICANT: Robinson, Murray O.  
TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: One Amgen Center Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951,733  
FILING DATE: 16-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/873,039  
FILING DATE: 11-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/751,189  
FILING DATE: 15-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Oleski, Nancy A.  
REGISTRATION NUMBER: 34,688  
REFERENCE/DOCKET NUMBER: A-433B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2848 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
SQ SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 100.0%; Score 2848; DB 22; Length 2848;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2848; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CACGGCTCCGGGAGCGCTGCTGCTGCGACAGTGGGAAGCCCTGGCCCCGGGACAC 60  
Qy 1 CACGGCTCCGGGAGCGCTGCGTCTGCTGCGACAGTGGGAAGCCCTGGCCCCGGGACAC 60

Db 61 CCCGCGATCCGGGCGCTCCCGCTGCCAGCGCTGGCGCTCCCTGCTGGCGAGCCACTAC 120  
Qy 61 CCCGCGATCCGGGCGCTCCCGCTGCCAGCGCTGGCGCTCCCTGCTGGCGAGCCACTAC 120

Db 121 CGCGAGGTGCTGCGGTGCGCACGTTCTGTCGGCGCTGGGCGCTGGGCGCTGGCGGTG 180  
Qy 121 CGCGAGGTGCTGCGGTGCGCACGTTCTGTCGGCGCTGGGCGCTGGGCGCTGGCGGTG 180

Db 181 GTGACGCGGGGACCCGGGCTTTCCGCGCGCTGGTGGCGCCAGTGCCTGTGTGCGTGS 240  
Qy 181 GTGACGCGGGGACCCGGGCTTTCCGCGCGCTGGTGGCGCCAGTGCCTGTGTGCGTGS 240

Db 241 CCTTGGAGCGAGCGCGCCCGCCCGCGCCCTCTTCCGCGAGGTGTCCTGCCTGAAG 300  
Qy 241 CCTTGGAGCGAGCGCGCGCCCGCCCGCGCCCTCTTCCGCGAGGTGTCCTGCCTGAAG 300

Db 301 GAGCTGTTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAAGACGTTGCTGGCC 360  
Qy 301 GAGCTGTTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGAAGACGTTGCTGGCC 360

Db 361 TTGCGCTTCGCGTGTGTCAGCGGGGCGCCCGGGGGCGCCCGCGAGGCTTACACACAGC 420  
Qy 361 TTGCGCTTCGCGTGTGTCAGCGGGGCGCCCGGGGGCGCCCGCGAGGCTTACACACAGC 420

Db 421 GTGCGAGCTACTGTCGCGCAACAGGTGACGAGCGCTGCGGGGAGCGGGCGTGGGG 480  
Qy 421 GTGCGAGCTACTGTCGCGCAACAGGTGACGAGCGCTGCGGGGAGCGGGCGTGGGG 480

Db 481 CTGCTGCTGCGCGCTGGCGGACAGCTGCTGTTTCACTGCTGTCAGCGCTGCGCGCTC 540  
Qy 481 CTGCTGCTGCGCGCTGGCGGACAGCTGCTGTTTCACTGCTGTCAGCGCTGCGCGCTC 540

Db 541 TTTGTGCTGGTGGCTCCAGCTGCGCTTACCAGGTGTGCGGCGCGCGCTGTACCAGCTC 600  
Qy 541 TTTGTGCTGGTGGCTCCAGCTGCGCTTACCAGGTGTGCGGCGCGCGCTGTACCAGCTC 600

Db 601 GCGCTGCGCACTAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660  
Qy 601 GCGCTGCGCACTAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

Db 661 TCGGAACCGGCTGGAACCATAGCTGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
Qy 661 TCGGAACCGGCTGGAACCATAGCTGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720

Db 721 CGGGTTCGAGGAGCGCGGGGCGAGTGCAGCGGAAGTCTGCCGTTGCCAAGAGGCC 780  
Qy 721 CGGGTTCGAGGAGCGCGGGGCGAGTGCAGCGGAAGTCTGCCGTTGCCAAGAGGCC 780

Db 781 AGCGTGGCGCTGCCCTGAGCGGAGCGGAGCGCGGTTGGGAGGGGCTTGGGCGCCAC 840  
Qy 781 AGCGTGGCGCTGCCCTGAGCGGAGCGGAGCGCGGTTGGGAGGGGCTTGGGCGCCAC 840

Db 841 CGGGGAGGACCGGTGACCGGAGTACCGTGTGTTCTGTGTGTGTCACCTGCCAGCCC 900  
Qy 841 CGGGGAGGACCGGTGACCGGAGTACCGTGTGTTCTGTGTGTGTCACCTGCCAGCCC 900

Db 901 GCCGAAGAAGCCACTCTTTTGAGGGTGGCTCTCTGCGACGCGCCACTTCCACCCATCC 960  
Qy 901 GCCGAAGAAGCCACTCTTTTGAGGGTGGCTCTCTGCGACGCGCCACTTCCACCCATCC 960

Db 961 GTGGGCGCGAGCAGCAGCGCGGCGCCCATCCACATCGCGGCGCCACAGCTCCCTGGGAC 1020  
Qy 961 GTGGGCGCGAGCAGCAGCGCGGCGCCCATCCACATCGCGGCGCCACAGCTCCCTGGGAC 1020

Db 1021 ACGCTTGTCCCGCGGTGTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAG 1080  
Qy 1021 ACGCTTGTCCCGCGGTGTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAG 1080

Db 1081 GAGCAGCTGCGGCGCTCTCTCTAGCTCTCTGAGGCGCCAGCTGACTGGGCTCGG 1140  
Qy 1081 GAGCAGCTGCGGCGCTCTCTCTAGCTCTCTGAGGCGCCAGCTGACTGGGCTCGG 1140

Db 1141 AGGCTGTTGAGAGCACTTTCTGCTGCTGAGGCGCTGAGTCCAGGAGCTTCCCGGAGG 1200  
Qy 1141 AGGCTGTTGAGAGCACTTTCTGCTGCTGAGGCGCTGAGTCCAGGAGCTTCCCGGAGG 1200

Db 1201 TTGCCCCGCTTCCCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
Qy 1201 TTGCCCCGCTTCCCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260

Db 1261 AACCAAGCGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
Qy 1261 AACCAAGCGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320

Db 1321 GTCACCCAGCAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
Qy 1321 GTCACCCAGCAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380

Db 1381 GAGGAGGAGGACAGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
Qy 1381 GAGGAGGAGGACAGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440

Db 1441 TGGCAGGTGTACGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
Qy 1441 TGGCAGGTGTACGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

Db 1501 GGCTCCAGGACAGAGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
Qy 1501 GGCTCCAGGACAGAGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560

Db 1561 AAGCATGCCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
Qy 1561 AAGCATGCCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620

Db 1621 TGGCTGCGCAGGAGCGCGGTGGCTGTGTTCCGCGCGCAGAGCAGCTGCTGCTGCTG 1680  
Qy 1621 TGGCTGCGCAGGAGCGCGGTGGCTGTGTTCCGCGCGCAGAGCAGCTGCTGCTGCTG 1680

Db 1681 GAGATCTTGGCCAAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
Qy 1681 GAGATCTTGGCCAAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

Db 1741 TCTTTCTTTTATGTACGAGACCACTTTTCAAGAAAGACAGCTCTTTTCTACCGGAAG 1800  
Qy 1741 TCTTTCTTTTATGTACGAGACCACTTTTCAAGAAAGACAGCTCTTTTCTACCGGAAG 1800

Db 1801 AGTGTCTGGAGCAAGTTCGAAAGCATTCGAAATCAGACAGCACTTGAAGAGGGTGCAGCTG 1860  
Qy 1801 AGTGTCTGGAGCAAGTTCGAAAGCATTCGAAATCAGACAGCACTTGAAGAGGGTGCAGCTG 1860

Db 1861 CGGGAGCTGTGGGAAGCAGAGTTCAGGAGCATTCGGGAAGCAGGCGCCCTGCTGAGC 1920  
Qy 1861 CGGGAGCTGTGGGAAGCAGAGTTCAGGAGCATTCGGGAAGCAGGCGCCCTGCTGAGC 1920

Db 1921 TCCAGACTCCGCTTCAATCCCAAGGCTGACGGCTGCGGGCTGCGGGCTGCGGGCTGCG 1980  
Qy 1921 TCCAGACTCCGCTTCAATCCCAAGGCTGACGGCTGCGGGCTGCGGGCTGCGGGCTGCG 1980



Db 1981 GTCTGGAGCCAGAGCTTCCGACAGAAAAGAGGCGCGAGCGTCTCACCTCAGGGTG 2040  
QY 1981 GTCTGGAGCCAGAGCTTCCGACAGAAAAGAGGCGCGAGCGTCTCACCTCAGGGTG 2040  
Db 2041 AAGGCACTGTTCAGCGTGTCAACTACGAGCGGCGCGGCGCCCTCTCTGGCGCC 2100  
QY 2041 AAGGCACTGTTCAGCGTGTCAACTACGAGCGGCGCGGCGCCCTCTCTGGCGCC 2100  
Db 2101 TCTGTGCTGGGCTGACGATATCCACAGGCGCTGGGCACTTGTGCTGCTGTGCGG 2160  
QY 2101 TCTGTGCTGGGCTGACGATATCCACAGGCGCTGGGCACTTGTGCTGCTGTGCGG 2160  
Db 2161 GCCAGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGTGAGCGCGGTAGCAC 2220  
QY 2161 GCCAGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGTGAGCGCGGTAGCAC 2220  
Db 2221 ACCATCCCCAGGACAGGCTCAGGAGGTTCATGCCAGCATCATCAAAACCCAGAACAG 2280  
QY 2221 ACCATCCCCAGGACAGGCTCAGGAGGTTCATGCCAGCATCATCAAAACCCAGAACAG 2280  
Db 2281 TACTGTGCTGCGTATGCGGTGTCAGAGGCGCCCATGGGCACTGTCGCAAGGCC 2340  
QY 2281 TACTGTGCTGCGTATGCGGTGTCAGAGGCGCCCATGGGCACTGTCGCAAGGCC 2340  
Db 2341 TTCAAGAGCCACGCTCTACTTTGACAGACTTCAGCGGTACATGCGACAGTTCGTGGCT 2400  
QY 2341 TTCAAGAGCCACGCTCTACTTTGACAGACTTCAGCGGTACATGCGACAGTTCGTGGCT 2400  
Db 2401 CACTGACGAGACCCCGCTGAGGATGCGCTCATCGACAGAGCTCTCCCTG 2460  
QY 2401 CACTGACGAGACCCCGCTGAGGATGCGCTCATCGACAGAGCTCTCCCTG 2460  
Db 2461 AATGAGCCAGAGTGGCTCTTCGAGCTTCTTAGCTTTCATGTGCGACAGCGCGT 2520  
QY 2461 AATGAGCCAGAGTGGCTCTTCGAGCTTCTTAGCTTTCATGTGCGACAGCGCGT 2520  
Db 2521 CGCATAGGGGCAAGTCTACGTCCTAGTCCAGGATCCGAGGATCCGAGGCTCCATCTCC 2580  
QY 2521 CGCATAGGGGCAAGTCTACGTCCTAGTCCAGGATCCGAGGATCCGAGGCTCCATCTCC 2580  
Db 2581 ACCTGCTCTGACGCTGTCTACGGGACATGAGAACAGCTGTTTGGGGGATTCGG 2640  
QY 2581 ACCTGCTCTGACGCTGTCTACGGGACATGAGAACAGCTGTTTGGGGGATTCGG 2640  
Db 2641 CGGACGGGCTGCTCTGCTTGGTGGATGATTTCTTGTGAGACCTCACCTCAC 2700  
QY 2641 CGGACGGGCTGCTCTGCTTGGTGGATGATTTCTTGTGAGACCTCACCTCAC 2700  
Db 2701 CACGCAAAACCTTCCTCAGGACCTTGTCCGAGGTGCTCCCTGAGTATGCTGCGTGG 2760  
QY 2701 CACGCAAAACCTTCCTCAGGACCTTGTCCGAGGTGCTCCCTGAGTATGCTGCGTGG 2760  
Db 2761 AACTTGGGAAGACAGTGTGAATCCCTGTAGAAGACGAGGCGCTTGGTGGCACGGCT 2820  
QY 2761 AACTTGGGAAGACAGTGTGAATCCCTGTAGAAGACGAGGCGCTTGGTGGCACGGCT 2820  
Db 2821 TTTGTTTCAGATGCGGCGCCACGGCTAT 2848  
QY 2821 TTTGTTTCAGATGCGGCGCCACGGCTAT 2848

RESULT 2  
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.  
AC xxxxxx  
DT

DE Sequence 19, Application US/08951733  
CC Sequence 19, Application US/08951733  
CC GENERAL INFORMATION:  
CC APPLICANT: Harrington, Lea A.  
CC APPLICANT: Robinson, Murray O.  
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
CC NUMBER OF SEQUENCES: 44  
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Amgen Inc.  
CC STREET: One Amgen Center Drive  
CC CITY: Thousand Oaks  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 91320-1789  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/951,733  
CC FILING DATE: 16-OCT-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/873,039  
CC FILING DATE: 11-JUN-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/751,189  
CC FILING DATE: 15-NOV-1996  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Oleski, Nancy A.  
CC REGISTRATION NUMBER: 34,688  
CC REFERENCE/DOCKET NUMBER: A-433B  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (805) 447-6504  
CC TELEFAX: (805) 499-8011  
CC INFORMATION FOR SEQ ID NO: 19:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3798 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
SQ SEQUENCE 3798 BP; 613 A; 1310 C; 1213 G; 662 T; 0 OTHER.  
  
Query Match 100.0%; Score 2848; DB 22; Length 3798;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2848; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 2 CACGCGTCCGGGACGCGCTCGCTCCCTGCTGCGACGCTGGGAAGCCCTGGCCCGGCCACC 61  
QY 1 CACGCGTCCGGGACGCGCTCGCTCCCTGCTGCGACGCTGGGAAGCCCTGGCCCGGCCACC 60  
Db 62 CCAGCGATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGACGCGCTAC 121  
QY 61 CCAGCGATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGACGCGCTAC 120  
Db 122 CCGAGGTGCTGCGCTGGCCACGTTGCTGCGGCGCTGGGGCCCGAGGGCTGGCGGCTG 181  
QY 121 CCGAGGTGCTGCGCTGGCCACGTTGCTGCGGCGCTGGGGCCCGAGGGCTGGCGGCTG 180  
Db 182 GTGACGCGGGGACCGCGGCTTTCCGCGCGCTGGGCGCCAGTCCCTGGTGTGCGTG 241  
QY 181 GTGACGCGGGGACCGCGGCTTTCCGCGCGCTGGTGGCCAGTCCCTGGTGTGCGTG 240  
Db 242 CCCTGGGACGACGCGCGCCCGCCCGCTCCCTGCTGCGCAGGTGCTGCTGCTGAAG 301  
QY 241 CCCTGGGACGACGCGCGCCCGCCCGCTCCCTGCTGCGCAGGTGCTGCTGCTGAAG 300  
Db 302 GAGCTGTGCGCGAGTGTGTCAGAGGCTGTGAGCGCGCGCGCGCGAGAACGCTTACACACAGC 361  
QY 301 GAGCTGTGCGCGAGTGTGTCAGAGGCTGTGAGCGCGCGCGCGCGAGAACGCTTACACACAGC 360  
Db 362 TTCGGCTTCGCGTGTGGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 421  
QY 361 TTCGGCTTCGCGTGTGGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
Db 422 GTGCGAGCTTACCTGCCCAACAGGTCAGCAGCGTGTGCGGCGGAGCGGGGCGTGGGG 481  
QY 421 GTGCGAGCTTACCTGCCCAACAGGTCAGCAGCGTGTGCGGCGGAGCGGGGCGTGGGG 480

Db 482 CTGCTGTCGCCCGGTGGCGACGACGTGCTGTTTCACTGCTGGCAGCGTGGCGGCTC 541  
QY |||||  
Db 481 CTGCTGTCGCCCGGTGGCGACGACGTGCTGTTTCACTGCTGGCAGCGTGGCGGCTC 540  
QY |||||  
Db 542 TTTGTGCTGGTGGTCTCCAGCTGGCGCTTACCAGGTGTGGCGGCGCGCTGTACCAAGCTC 601  
QY |||||  
Db 541 TTTGTGCTGGTGGTCTCCAGCTGGCGCTTACCAGGTGTGGCGGCGCGCTGTACCAAGCTC 600  
QY |||||  
Db 602 GGGCGTGCCACTAGCGCCCGCGCCCGCCACACGCTAGTGTAGCCCGCAAGGCGTCTGGGA 661  
QY |||||  
Db 601 GGGCGTGCCACTAGCGCCCGCGCCCGCCACACGCTAGTGTAGCCCGCAAGGCGTCTGGGA 660  
QY |||||  
Db 662 TGCGAAGCGGCTTGAACCATAGCTCAGGAGCGCGGGGTCCCGCTGGGCTTCCAGCC 721  
QY |||||  
Db 661 TGCGAAGCGGCTTGAACCATAGCTCAGGAGCGCGGGGTCCCGCTGGGCTTCCAGCC 720  
QY |||||  
Db 722 CCGGTGCGAGAGCGCGGGGGAGTGCCAGCGGAAGTCTGCGGTGTGCCCAAGAGGCC 781  
QY |||||  
Db 721 CCGGTGCGAGAGCGCGGGGGAGTGCCAGCGGAAGTCTGCGGTGTGCCCAAGAGGCC 780  
QY |||||  
Db 782 AGGCGTGCGGTGCCCTTGAGCCGGAGCGGACGCCCGTGTGGGCAGGGGTCTTGGGCCAC 841  
QY |||||  
Db 781 AGGCGTGCGGTGCCCTTGAGCCGGAGCGGACGCCCGTGTGGGCAGGGGTCTTGGGCCAC 840  
QY |||||  
Db 842 CCGGCAAGAGCCACTCTTTTGAGGGTGGCTCTGTGGCAGCGGCACTCCACCCATCC 961  
QY |||||  
Db 901 GCCGAAGAGCCACTCTTTTGAGGGTGGCTCTGTGGCAGCGGCACTCCACCCATCC 960  
QY |||||  
Db 962 GTGGGCGCCAGCACAGCGGGGCCCCCATCCACATCGCGGCGCCACACGCTTCCCTGGGAC 1021  
QY |||||  
Db 961 GTGGGCGCCAGCACAGCGGGGCCCCCATCCACATCGCGGCGCCACACGCTTCCCTGGGAC 1020  
QY |||||  
Db 1022 ACGCTTGTCCCGGTGTAGCCGAGACCAAGCACATTCCTTCTACTCTCAGGCGACAAG 1081  
QY |||||  
Db 1021 ACGCTTGTCCCGGTGTAGCCGAGACCAAGCACATTCCTTCTACTCTCAGGCGACAAG 1080  
QY |||||  
Db 1082 GAGCAGTGTGGCGCTCTTCTACTAGCTCTCTGAGGCGCCAGCTGACTTGGCGCTCGG 1141  
QY |||||  
Db 1081 GAGCAGTGTGGCGCTCTTCTACTAGCTCTCTGAGGCGCCAGCTGACTTGGCGCTCGG 1140  
QY |||||  
Db 1142 AGGCTGTGAGACCATCTTTTCTGGGTTCAGGCCCTGGATGCCAGGACTTCCCGGAGG 1201  
QY |||||  
Db 1141 AGGCTGTGAGACCATCTTTTCTGGGTTCAGGCCCTGGATGCCAGGACTTCCCGGAGG 1200  
QY |||||  
Db 1202 TTGCCCGGCTTGGCCAGCGCTACTGGCAATGGGCCCTGTTTCTGGAGCTGCTTGGG 1261  
QY |||||  
Db 1201 TTGCCCGGCTTGGCCAGCGCTACTGGCAATGGGCCCTGTTTCTGGAGCTGCTTGGG 1260  
QY |||||  
Db 1262 AACACGCGCAGTGGCCCTACGGGGTGTCTCTAAGACGCACTGCCGCTGCGAGCTGGG 1321  
QY |||||  
Db 1261 AACACGCGCAGTGGCCCTACGGGGTGTCTCTAAGACGCACTGCCGCTGCGAGCTGGG 1320  
QY |||||  
Db 1322 GTCACCCAGCAGCGGTGTGTGTGCCGGGAGAACGCCAGGCTGTGTGGCGGCCGCC 1381  
QY |||||  
Db 1321 GTCACCCAGCAGCGGTGTGTGTGCCGGGAGAACGCCAGGCTGTGTGGCGGCCGCC 1380  
QY |||||  
Db 1382 GAGGAGGAGACAGACCCCGCTGCTGTGTGAGCTGTCCGCCAGCAGCAGGCC 1441  
QY |||||  
Db 1381 GAGGAGGAGACAGACCCCGCTGCTGTGTGAGCTGTCCGCCAGCAGCAGGCC 1440  
QY |||||  
Db 1442 TGGCAGGTGTACGGCTTGTGTGGGGCTGTCTGCGCGGCTGTGTGCCCGCAGGCTCTGG 1501  
QY |||||  
Db 1441 TGGCAGGTGTACGGCTTGTGTGGGGCTGTCTGCGCGGCTGTGTGCCCGCAGGCTCTGG 1500  
QY |||||  
Db 1502 GGCTCCAGGCACACGAACCGGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGG 1561  
QY |||||  
Db 1501 GGCTCCAGGCACACGAACCGGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGG 1560  
QY |||||  
Db 1562 AAGCATGCCAAGCTCTCGCTCCAGAGCTGACGTGGAGATGACGGTGGCGGACTGGCT 1621

QY |||||  
Db 1561 AAGCATGCCAAGCTCTCGCTCCAGAGCTGACGTGGAAGATGAGCGTCCGGGACTGGCT 1620  
QY |||||  
Db 1622 TGGCTGCGCAGGAGCCAGGGGTGGCTGTGTTCCGCCCGCAGAGCACCGTCTCCGTGAG 1681  
QY |||||  
Db 1621 TGGCTGCGCAGGAGCCAGGGGTGGCTGTGTTCCGCCCGCAGAGCACCGTCTCCGTGAG 1680  
QY |||||  
Db 1682 GAGATCTGGCCAAAGTCTTCGCACTGGCTGATGAGTGTGTACGTGTGAGCTGTCTGAG 1741  
QY |||||  
Db 1681 GAGATCTGGCCAAAGTCTTCGCACTGGCTGATGAGTGTGTACGTGTGAGCTGTCTGAG 1740  
QY |||||  
Db 1742 TCTTCTTTTATCTCAGGAGACCGCTTCAAAGAACAGGCTCTTTTCTACCGGAAG 1801  
QY |||||  
Db 1741 TCTTCTTTTATCTCAGGAGACCGCTTCAAAGAACAGGCTCTTTTCTACCGGAAG 1800  
QY |||||  
Db 1802 AGTGTCTGGAGCAAGTGTGAAGCATTTGAATCAGACAGCACTTGAAGAGGGTGCAGCTG 1861  
QY |||||  
Db 1801 AGTGTCTGGAGCAAGTGTGAAGCATTTGAATCAGACAGCACTTGAAGAGGGTGCAGCTG 1860  
QY |||||  
Db 1862 CGGAGCTGTGGAAGCAGAGGTCTAGGAGCATCGGGAAGCCAGGCCCGCTGTGAGC 1921  
QY |||||  
Db 1861 CGGAGCTGTGGAAGCAGAGGTCTAGGAGCATCGGGAAGCCAGGCCCGCTGTGAGC 1920  
QY |||||  
Db 1922 TCCAGACTCCGCTTCATCCCAAGCCTGAGGGCTGCGGGCTGTGAACATGGACTAC 1981  
QY |||||  
Db 1921 TCCAGACTCCGCTTCATCCCAAGCCTGAGGGCTGCGGGCTGTGAACATGGACTAC 1980  
QY |||||  
Db 1982 GTGCTGGGAGCCAGAGCTTCCCGAGAAAAGAGGCGCGAGCGTCTCACTCGAGGGTG 2041  
QY |||||  
Db 1981 GTGCTGGGAGCCAGAGCTTCCCGAGAAAAGAGGCGCGAGCGTCTCACTCGAGGGTG 2040  
QY |||||  
Db 2042 AAGGCACTGTTCAGCTGTCTAACTACGAGCGGCGCGCGCGCCCGCTTCTGGGCGCC 2101  
QY |||||  
Db 2041 AAGGCACTGTTCAGCTGTCTAACTACGAGCGGCGCGCGCGCCCGCTTCTGGGCGCC 2100  
QY |||||  
Db 2102 TCTGTCTGGGCTTGACGATATCCACAGGCGCTGCGGACCTTCTGCTGCGTGTGCGG 2161  
QY |||||  
Db 2101 TCTGTCTGGGCTTGACGATATCCACAGGCGCTGCGGACCTTCTGCTGCGTGTGCGG 2160  
QY |||||  
Db 2162 GCCAGAGCCCGCGCTGAGCTGTACTTTGCAAGTGTGATGTGAGCGGCGCTGACGAC 2221  
QY |||||  
Db 2161 GCCAGAGCCCGCGCTGAGCTGTACTTTGCAAGTGTGATGTGAGCGGCGCTGACGAC 2220  
QY |||||  
Db 2222 ACCATCCCGCAGGACAGGCTCAGGAGGTATCGCAGCATCATCAACCCCAAGACAG 2281  
QY |||||  
Db 2221 ACCATCCCGCAGGACAGGCTCAGGAGGTATCGCAGCATCATCAACCCCAAGACAG 2280  
QY |||||  
Db 2282 TACTGGCTGCTGCTGATCCGCTGCTGAGAGCGGCGCCATGGGACGCTCCGCAAGGCC 2341  
QY |||||  
Db 2281 TACTGGCTGCTGCTGATCCGCTGCTGAGAGCGGCGCCATGGGACGCTCCGCAAGGCC 2340  
QY |||||  
Db 2342 TTCAAGAGCCACGCTCTCTACTTTGACAGACCTCCAGCCGTACATGGACAGTTCGTGGCT 2401  
QY |||||  
Db 2341 TTCAAGAGCCACGCTCTCTACTTTGACAGACCTCCAGCCGTACATGGACAGTTCGTGGCT 2400  
QY |||||  
Db 2402 CACCTGCGAGGAGACAGCCGCTGAGGGATGCGTGTATCATGAGCAGAGCTCCTCCCTG 2461  
QY |||||  
Db 2401 CACCTGCGAGGAGACAGCCGCTGAGGGATGCGTGTATCATGAGCAGAGCTCCTCCCTG 2460  
QY |||||  
Db 2462 AATGAGCCAGCAGTGGCTTCTTCGAGCTTCTTCTACGCTTCTATGTCACACAGCGGTG 2521  
QY |||||  
Db 2461 AATGAGCCAGCAGTGGCTTCTTCGAGCTTCTTCTACGCTTCTATGTCACACAGCGGTG 2520  
QY |||||  
Db 2522 CGCATAGGGGCAAGTCTCTACGTCCAGTGCAGGGATCCCGCAGGCGTCCATCTCTCC 2581  
QY |||||  
Db 2521 CGCATAGGGGCAAGTCTCTACGTCCAGTGCAGGGATCCCGCAGGCGTCCATCTCTCC 2580  
QY |||||  
Db 2582 ACCTGCTGTGACGCTGTGCTACGCGACATGGAACAAAGCTGTTTGGGGGATTCGG 2641  
QY |||||  
Db 2581 ACCTGCTGTGACGCTGTGCTACGCGACATGGAACAAAGCTGTTTGGGGGATTCGG 2640  
QY |||||  
Db 2642 CGGAGCGGCTGCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCCTACCC 2701  
QY |||||

QY 2641 CGGACGGGCTCTCTGCTGTTGGTGATGATTTCTTTGTTGGTACACCTCACCTCACC 2700  
Db 2702 CACCGAAACCTCTCTCAGACCTGGTCGGAGGTGTCCTTGATGATGGCTGGGTGGTG 2761  
QY 2701 CACCGAAACCTCTCTCAGACCTGGTCGGAGGTGTCCTTGATGATGGCTGGGTGGTG 2760  
Db 2762 AACTTGGGGAAGCAGTGGTGAACCTCCCTGTAGAAGCAGGCGCTGGGTGGGACGGCT 2821  
QY 2761 AACTTGGGGAAGCAGTGGTGAACCTCCCTGTAGAAGCAGGCGCTGGGTGGGACGGCT 2820  
Db 2822 TTGTTTCAGATGCGGCGCCACGGCCTAT 2849  
QY 2821 TTGTTTCAGATGCGGCGCCACGGCCTAT 2848

RESULT 3  
ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.  
AC xxxxxx

Sequence 1, Application US/08911312  
Sequence 1, Application US/08911312  
GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911.312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-002500US  
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 7029 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 782..4177  
CC OTHER INFORMATION: /product= "human telomerase reverse  
CC OTHER INFORMATION: transcriptase (hprt)"  
CC OTHER INFORMATION: /note= "cDNA contained in plasmid  
CC OTHER INFORMATION: pGRN121"  
CC SEQUENCE 7029 BP; 1416 A; 2122 C; 2051 G; 1440 T; 0 OTHER.  
Query Match 99.78; Score 2840; DB 22; Length 7029;  
Best Local Similarity 99.98; Pred. No. 0.00e+00;  
Matches 2843; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Db 718 CGCGTGCAGCAGCGCTGCTGCTGCACGTGGGAAGCCTGGCCCGGCCACCCC 777  
QY 3 CGGTCCGGCAGCGCTGCTGCTGCGCACGTGGGAAGCCTGGCCCGGCCACCCC 62  
Db 778 CGCGATGCCCGCGCTCCCGCTGCCGAGCGTGGCGCTCCCTGCTGCGCAGCCACTACCG 837  
QY 63 CGCGATGCCCGCGCTCCCGCTGCCGAGCGTGGCGCTCCCTGCTGCGCAGCCACTACCG 122  
Db 838 CGAGGTGCTCCGCTGGCGCAGCTTGTGCGGCGCTGGGCGCCCGAGGCGTGGGCGCTGT 897  
QY 123 CGAGGTGCTCCGCTGGCGCAGCTTGTGCGGCGCTGGGCGCCCGAGGCGTGGGCGCTGT 182  
Db 898 GCAGCGGGGACCGCGCGCTTTCCGCGCGTGGTGGCGCCAGTGCCTGCTGTCGTCGCC 957  
QY 183 GCAGCGGGGACCGCGCGCTTTCCGCGCGTGGTGGCGCCAGTGCCTGCTGTCGTCGCC 242  
Db 958 CTGGGACGACGGCG 1017  
QY 243 CTGGGACGACGGCG 302  
Db 1018 GCTGGTGGCGCGAGTGTGTCAGAGGCTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1077  
QY 303 GCTGGTGGCGCGAGTGTGTCAGAGGCTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 362  
Db 1078 CGGCTTCGCGCTGCTGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137  
QY 363 CGGCTTCGCGCTGCTGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 422  
Db 1138 GCGCAGCTACTGCCCCAACACGCGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGCT 1197  
QY 423 GCGCAGCTACTGCCCCAACACGCGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGCT 482  
Db 1198 GCTGCTGCGCGCGTGGGCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257  
QY 483 GCTGCTGCGCGCGTGGGCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542  
Db 1258 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317  
QY 543 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602  
Db 1318 CGCTGCCACTCAGGCG 1377  
QY 603 CGCTGCCACTCAGGCG 662  
Db 1378 CGAACGGGCTGGAACCATAGCTAGGAGCGCGGGGTCCCGCTGGGCGCTGCCAGCCCC 1437  
QY 663 CGAACGGGCTGGAACCATAGCTAGGAGCGCGGGGTCCCGCTGGGCGCTGCCAGCCCC 722  
Db 1438 GGTGTCGAGGAGCGCGGGGCGAGTGCCACCGCCAGTCTGCCGTTGCCCAAGAGGCCAG 1497  
QY 723 GGTGTCGAGGAGCGCGGGGCGAGTGCCACCGCCAGTCTGCCGTTGCCCAAGAGGCCAG 782

Db 1498 GCGTGGCGCTGCCCTTGAGCCGGAGCGAGCCCGTTGGGAGGGGTCCTTGGGGCCACCC 1557  
Qy 783 GCGTGGCGCTGCCCTTGAGCCGGAGCGAGCCCGTTGGGAGGGGTCCTTGGGGCCACCC 842  
Db 1558 GGGCAGGACGGGTGACCGAGTGCCTGTTCTGCTGTGTGTACCTGCGACCGCGC 1617  
Qy 843 GGGCAGGACGGGTGACCGAGTGCCTGTTCTGCTGTGTGTACCTGCGACCGCGC 902  
Db 1618 CGAAGAACCCACCTTTTGGAGGTGCGCTCTCTGGCAGCGGCCACTCCACCCATCCGT 1677  
Qy 903 CGAAGAACCCACCTTTTGGAGGTGCGCTCTCTGGCAGCGGCCACTCCACCCATCCGT 962  
Db 1678 GGGCGCCAGACACGCGGGGCCCCCAATCCACATCGCGGCCACACGTCCTCGGACAC 1737  
Qy 963 GGGCGCCAGACACGCGGGGCCCCCAATCCACATCGCGGCCACACGTCCTCGGACAC 1022  
Db 1738 GCCTTGTCCCGGTTGACGCGAGCAACGACTTCTCTACTCTCAGCGGCACAAAGGA 1797  
Qy 1023 GCCTTGTCCCGGTTGACGCGAGCAACGACTTCTCTACTCTCAGCGGCACAAAGGA 1082  
Db 1798 CGAGTGGCGGCCCTCTCTACTCAGCTCTCTGAGGCCCAAGCCTGACTGCGCTCGGAG 1857  
Qy 1083 CGAGTGGCGGCCCTCTCTACTCAGCTCTCTGAGGCCCAAGCCTGACTGCGCTCGGAG 1142  
Db 1858 GCTGCTGAGACCATCTTCTGCGTTGAGGCTCAGGCCCTGAGTGCAGGGACTCCCGGAGTT 1917  
Qy 1143 GCTGCTGAGACCATCTTCTGCGTTGAGGCTCAGGCCCTGAGTGCAGGGACTCCCGGAGTT 1202  
Db 1918 GCGCCGCTGCGCCACGCGCTACTGCAAAATCGCGCCCTGTTCTTGGAGTGTGTTGGAA 1977  
Qy 1203 GCGCCGCTGCGCCACGCGCTACTGCAAAATCGCGCCCTGTTCTTGGAGTGTGTTGGAA 1262  
Db 1978 CCACGCGAGTGCCTTACGGGGTCTCTCAAGACGACACTGCCGCTGCGAGTGCCTG 2037  
Qy 1263 CCACGCGAGTGCCTTACGGGGTCTCTCAAGACGACACTGCCGCTGCGAGTGCCTG 1322  
Db 2038 CACCCACAGCGGTGCTGTGCCGGGAGAGGCCCCAGGGCTGTGTGGGGGCCCGGA 2097  
Qy 1323 CACCCACAGCGGTGCTGTGCCGGGAGAGGCCCCAGGGCTGTGTGGGGGCCCGGA 1382  
Db 2098 GGAGGAGACACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157  
Qy 1383 GGAGGAGACACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1442  
Db 2158 CGAGGTGACGGCTTCTGTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2217  
Qy 1443 CGAGGTGACGGCTTCTGTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1502  
Db 2218 CTCAGGACACAGACCCCGCTTCTCAGGAACACAGAGTTCATCTCCTCGGGGA 2277  
Qy 1503 CTCAGGACACAGACCCCGCTTCTCAGGAACACAGAGTTCATCTCCTCGGGGA 1562  
Db 2278 GCATGCCAAGCTCTCGCTGACAGGCTCAGCTGGAAGATGACGTCGCGGACTGCGCTTG 2337  
Qy 1563 GCATGCCAAGCTCTCGCTGACAGGCTCAGCTGGAAGATGACGTCGCGGACTGCGCTTG 1622  
Db 2338 GCTGCGAGGAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGACACCTGCTGCGTGAGGA 2397  
Qy 1623 GCTGCGAGGAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGACACCTGCTGCGTGAGGA 1682  
Db 2398 GATCTGCGCAAGTTCCTGCACTGCTGATGATGATGATGATGATGATGATGATGATGAT 2457  
Qy 1683 GATCTGCGCAAGTTCCTGCACTGCTGATGATGATGATGATGATGATGATGATGATGAT 1742  
Db 2458 TTCTCTTTTATGTACGGAGACACCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAG 2517  
Qy 1743 TTCTCTTTTATGTACGGAGACACCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAG 1802  
Db 2518 TGTCTGGAGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGACGCTCGC 2577  
Qy 1803 TGTCTGGAGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGACGCTCGC 1862

Db 2578 GGAGCTCTCGGAAGCAGAGTCAAGCAGCATCGGAAGCCAGGCCGCCCTGCTGACGTC 2637  
Qy 1863 GGAGCTCTCGGAAGCAGAGTCAAGCAGCATCGGAAGCCAGGCCGCCCTGCTGACGTC 1922  
Db 2638 CAGACTCCGCTTCAATCCCAAGCTGACGGCTGCGGCCGATTTGTGAACATGACTACGT 2697  
Qy 1923 CAGACTCCGCTTCAATCCCAAGCTGACGGCTGCGGCCGATTTGTGAACATGACTACGT 1982  
Db 2698 CGTGGGAGCCAGAACCTTCCGAGAGAAAAGAGGCCGAGCGCTCTCACTCGAGGGTGAA 2757  
Qy 1983 CGTGGGAGCCAGAACCTTCCGAGAGAAAAGAGGCCGAGCGCTCTCACTCGAGGGTGAA 2042  
Db 2758 GGCACCTGTTGAGCGTGTCAACTACAGCGGGCGCGGCCGCCCTCTTGGGGCGCTC 2817  
Qy 2043 GGCACCTGTTGAGCGTGTCAACTACAGCGGGCGCGGCCGCCCTCTTGGGGCGCTC 2102  
Db 2818 TGTGCTGGGCTGACGATATCACAGGGCTGCGCACCTTCTGCTGTGCTGTGCGGGC 2877  
Qy 2103 TGTGCTGGGCTGACGATATCACAGGGCTGCGCACCTTCTGCTGTGCTGTGCGGGC 2162  
Db 2878 CCAGGACCCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGAC 2937  
Qy 2163 CCAGGACCCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGAC 2222  
Db 2938 CATCCCCAGACAGGCTCAGGAGTATCGCCAGCATCATCAAAACCCAGAACAGTA 2997  
Qy 2223 CATCCCCAGACAGGCTCAGGAGTATCGCCAGCATCATCAAAACCCAGAACAGTA 2282  
Db 2998 CTGCGTCCGCTGCGGTATGCCGTGTCCAGAAAGCGCCCATGGGACGCTCCGCAAGGCTT 3057  
Qy 2283 CTGCGTCCGCTGCGGTATGCCGTGTCCAGAAAGCGCCCATGGGACGCTCCGCAAGGCTT 2342  
Db 3058 CAAGAGCCACGCTCTACTTGTGACAGACCTCCAGCCGATACATGCGACAGTTCGTTGGTCA 3117  
Qy 2343 CAAGAGCCACGCTCTACTTGTGACAGACCTCCAGCCGATACATGCGACAGTTCGTTGGTCA 2402  
Db 3118 CTGCGAGACACAGCCCGCTGAGGATGCGCTGCTATCGAGCAGAGTCTCTCCCTGAA 3177  
Qy 2403 CTGCGAGACACAGCCCGCTGAGGATGCGCTGCTATCGAGCAGAGTCTCTCCCTGAA 2462  
Db 3178 TGAGGCCAGAGTGGCGCTTTCGACGCTTCTACGCTTTCATGTGCCACACGCGGTGCG 3237  
Qy 2463 TGAGGCCAGAGTGGCGCTTTCGACGCTTCTACGCTTTCATGTGCCACACGCGGTGCG 2522  
Db 3238 CATCAGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGAGGGTCCATCTCTCTCCAC 3297  
Qy 2523 CATCAGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGAGGGTCCATCTCTCTCCAC 2582  
Db 3298 GCTGCTCTGACGCTGTGCTACGGCGACATGGAGAACAGCTGTTTGGGGGATTCGGCG 3357  
Qy 2583 GCTGCTCTGACGCTGTGCTACGGCGACATGGAGAACAGCTGTTTGGGGGATTCGGCG 2642  
Db 3358 GGAGCGGCTGCTCTGCGTTTGGTGGATGATTTCTTCTGTTGACACCTCAGCTCACCAC 3417  
Qy 2643 GGAGCGGCTGCTCTGCGTTTGGTGGATGATTTCTTCTGTTGACACCTCAGCTCACCAC 2702  
Db 3418 CGGAAAACCTTCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3477  
Qy 2703 CGGAAAACCTTCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2762  
Db 3478 CTGCGGAAGACAGTGTGAACCTTCCCTGTAGAAAGACGAGGCCCTTGGGTGGCAGGCTTT 3537  
Qy 2763 CTGCGGAAGACAGTGTGAACCTTCCCTGTAGAAAGACGAGGCCCTTGGGTGGCAGGCTTT 2822  
Db 3538 TGTTCAGATCGCGGCCACGCGCTAT 3563  
Qy 2823 TGTTCAGATCGCGGCCACGCGCTAT 2848

RESULT 4

ID US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.

AC XXXXX

DT









QY 1392 CACAGACCCCGTCCGCTGGTGCAGCTGCTCGGCCAGACAGCCCGCTGGCAGGTGTA 1451  
Db 1441 CGGCTTGTGGGGCTGCTGCGCGCGCTGGTGGCCCCAGCCCTCTGGGGCTCCAGGCA 1500  
QY 1452 CGGCTTGTGGGGCTGCTGCGCGCGCTGGTGGCCCCAGCCCTCTGGGGCTCCAGGCA 1511  
Db 1501 CAACGAAGCCGCTTCCCTCAGAGACACCAAGAAGTTATCTCCCTGGGGAAGCATGCCAA 1560  
QY 1512 CAACGAAGCCGCTTCCCTCAGAGACACCAAGAAGTTATCTCCCTGGGGAAGCATGCCAA 1571  
Db 1561 GCTCTCGCTCAGGAGCTGAGCTGGAAGATGAGCGTGGGGACTGCGCTTGGCTGCGGAG 1620  
QY 1572 GCTCTCGCTCAGGAGCTGAGCTGGAAGATGAGCGTGGGGACTGCGCTTGGCTGCGGAG 1631  
Db 1621 GAGCCAGGGTGGCTGTGTTCGGCGCCGACAGACCGCTGTGCTGAGGAGATCTCTGGC 1680  
QY 1632 GAGCCAGGGTGGCTGTGTTCGGCGCCGACAGACCGCTGTGCTGAGGAGATCTCTGGC 1691  
Db 1681 CAAGTTCTGCACTGGCTGATGATGTGTAGTGTGCTGAGCTGCTCAGGTCTTCTTTTA 1740  
QY 1692 CAAGTTCTGCACTGGCTGATGATGTGTAGTGTGCTGAGCTGCTCAGGTCTTCTTTTA 1751  
Db 1741 TGTCAAGGAGACCACTGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1800  
QY 1752 TGTCAAGGAGACCACTGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1811  
Db 1801 CAAGTTGCAAGCAATGGAATCAGACAGCACTTGAAGAGGTGACGTGCGGGAGCTGTC 1860  
QY 1812 CAAGTTGCAAGCAATGGAATCAGACAGCACTTGAAGAGGTGACGTGCGGGAGCTGTC 1871  
Db 1861 GGAAGCAGAGTCAAGCAGATCGGAGAGCCAGGCCCGCTGCTGAGCTCCAGACTCCG 1920  
QY 1872 GGAAGCAGAGTCAAGCAGATCGGAGAGCCAGGCCCGCTGCTGAGCTCCAGACTCCG 1931  
Db 1921 CTTTCATCCCAAGCTGACGGCTGCGCGCATTTGTGAACATGACTACGTCCTGGGAGC 1980  
QY 1932 CTTTCATCCCAAGCTGACGGCTGCGCGCATTTGTGAACATGACTACGTCCTGGGAGC 1991  
Db 1981 CAGAACGTTCCGCGAGAAAAGAGGGCGGCGCTCTACCTCGAGGGTGAAGGCACTGTT 2040  
QY 1992 CAGAACGTTCCGCGAGAAAAGAGGGCGGCGCTCTACCTCGAGGGTGAAGGCACTGTT 2051  
Db 2041 CAGCGTGTCAACTACGAGCGGGCGGCGCGCGCGCTCTGCGCGCTCTGCTGTGGG 2100  
QY 2052 CAGCGTGTCAACTACGAGCGGGCGGCGCGCGCGCTCTGCGCGCTCTGCTGTGGG 2111  
Db 2101 CCGTGCAGATATCACAGGGCTGCGCGCACCTTGTGTGCTGCTGCGGGCCAGGACCC 2160  
QY 2112 CCGTGCAGATATCACAGGGCTGCGCGCACCTTGTGTGCTGCTGCGGGCCAGGACCC 2171  
Db 2161 GCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2220  
QY 2172 GCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCCCCA 2231  
Db 2221 GGAAGGCTCAGGAGTATCGGAGCATATCAAAACCCAGAACACGATGCTGCTGGG 2280  
QY 2232 GGAAGGCTCAGGAGTATCGGAGCATATCAAAACCCAGAACACGATGCTGCTGGG 2291  
Db 2281 TCGGTATCGGTGTCAGAGAGCGCCCATGCGGACGCTCGCGAAGGCTTCAAGAGCCA 2340  
QY 2292 TCGGTATCGGTGTCAGAGAGCGCCCATGCGGACGCTCGCGAAGGCTTCAAGAGCCA 2351  
Db 2341 CGTCTCTACCTTGACAGACCTCCAGCGCTACATGCGACAGTTCGTGGCTACCTGCGAGA 2400  
QY 2352 CGTCTCTACCTTGACAGACCTCCAGCGCTACATGCGACAGTTCGTGGCTACCTGCGAGA 2411  
Db 2401 GACAGCCCGCTGAGGATGCGCTGCTATCGAGCAGAGTCTCTCCCTGATGAGGCGAG 2460  
QY 2412 GACAGCCCGCTGAGGATGCGCTGCTATCGAGCAGAGTCTCTCCCTGATGAGGCGAG 2471  
Db 2461 CAGTGGCTCTTTCGACGCTTCTTCCAGCTTCTATGTGCCACACCGCTGCGCATCAGGG 2520  
QY 2472 CAGTGGCTCTTTCGACGCTTCTTCCAGCTTCTATGTGCCACACCGCTGCGCATCAGGG 2531

## RESULT 6

ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.

AC xxxxxx

DT

DE Sequence 1, Application US/09052919

CC Sequence 1, Application US/09052919

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: Antisense Compositions for Detecting and

CC TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 72

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94111-3834

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/09/052,919

CC FILING DATE: 31-MAR-1998

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/854,050







CC	LENGTH: 4023 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
SQ	SEQUENCE 4023 BP; 668 A; 1363 C; 1277 G; 715 T; 0 OTHER.	
Query Match 99.5%; Score 2834; DB 23; Length 4023; Best Local Similarity 99.98; Pred. No. 0.00e+00; Matches 2836; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Db	3 GGCAGCGCTCGCTCTGCTGCGACAGTGGGAAGCCCTTGGCCCGCCGACACCCCGCGATGC 62	1022
QY		
Db	11 GGCAGCGCTCGCTCTGCTGCGACAGTGGGAAGCCCTTGGCCCGCCGACACCCCGCGATGC 70	1030
QY		
Db	63 CGCGCTCCCGTCCGAGCGTGGCTCCCTGCTGCGAGCACTACGCGAGGTGC 122	1082
QY		
Db	71 CGCGCTCCCGTCCGAGCGTGGCTCCCTGCTGCGAGCACTACGCGAGGTGC 130	1090
QY		
Db	123 TGCGCTGGCACGTTCTGCGCGCTTGGGCCCGCCAGGCTGGCGTGGCAGCGC 182	1142
QY		
Db	131 TGCGCTGGCACGTTCTGCGCGCTTGGGCCCGCCAGGCTGGCGTGGCAGCGC 190	1150
QY		
Db	183 GGGACCGCGCGCTTCCGCGCGTGGTGGCCCAAGTGCCTGGTGGCTGGCGACG 242	1202
QY		
Db	191 GGGACCGCGCGCTTCCGCGCGTGGTGGCCCAAGTGCCTGGTGGCTGGCGACG 250	1210
QY		
Db	243 CAGCGCGCGCGCGCGCGCTTCCCTTCCGCGCAGGTGCTGCGTGAAGAGCTGTGG 302	1262
QY		
Db	251 CAGCGCGCGCGCGCGCGCTTCCCTTCCGCGCAGGTGCTGCTGAAGAGCTGTGG 310	1270
QY		
Db	303 CCGAGTGTGACAGAGCTGTGCGAGCGCGCGCGAGAGCTGCTGGCCTTGGCTTCG 362	1322
QY		
Db	311 CCGAGTGTGACAGAGCTGTGCGAGCGCGCGCGAGAGCTGCTGGCCTTGGCTTCG 370	1330
QY		
Db	363 CGCTGTGTGAGCGCGCGCGCGCGCGCTTCCAGAGCGCTTCCAGCAGCGTGGCGAGCT 422	1382
QY		
Db	371 CGCTGTGTGAGCGCGCGCGCGCGCGCTTCCAGAGCGCTTCCAGCAGCGTGGCGAGCT 430	1390
QY		
Db	423 ACCTGCCCAACACGGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGCGTGTGTC 482	1442
QY		
Db	431 ACCTGCCCAACACGGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGCGTGTGTC 490	1450
QY		
Db	483 GC CGCGTGGCGAGAGCTGCTGCTTACCTGTGCGACGCTGGCGCTCTTTGTGTCG 542	1502
QY		
Db	491 GC CGCGTGGCGAGAGCTGCTGCTTACCTGTGCGACGCTGGCGCTCTTTGTGTCG 550	1510
QY		
Db	543 TGGCTCCAGCTGGCTTACCAGTGTGCGGGCGCGCTGTACAGCTGGCGCTGCCA 602	1562
QY		
Db	551 TGGCTCCAGCTGGCTTACCAGTGTGCGGGCGCGCTGTACAGCTGGCGCTGCCA 610	1570
QY		
Db	603 CTCAGGCCCGCGCGCGCGCGCTAGTGGACCCCGAAGCGCTGGGATGCGAAGCGG 662	1622
QY		
Db	611 CTCAGGCCCGCGCGCGCGCTAGTGGACCCCGAAGCGCTGGGATGCGAAGCGG 670	1630
QY		
Db	663 CCTGGACCACTAGCTCAGGAGCGCGGCTTCCCTTGGCGCTGCCAGCCCGGGTGC 722	1682
QY		
Db	671 CCTGGACCACTAGCTCAGGAGCGCGGCTTCCCTTGGCGCTGCCAGCCCGGGTGC 730	1690
QY		
Db	723 GGAGCGCGGGGCGAGTGCCAGCGAAGTCTGCGTTGCCAAGAGCGCCAGGCGTGGCG 782	1742
QY		
Db	731 GGAGCGCGGGGCGAGTGCCAGCGAAGTCTGCGTTGCCAAGAGCGCCAGGCGTGGCG 790	1750
QY		
Db	783 CTGCCCCGTGAGCGGAGCGGACGCCCTTGGGCAAGGGTCTTGGGCCACCCCGGCGAG 842	1802
QY		
Db	791 CTGCCCCGTGAGCGGAGCGGACGCCCTTGGGCAAGGGTCTTGGGCCACCCCGGCGAG 850	1810
QY		
Db	843 CGCGTGGACCGAGTGACGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902	1862
QY		
Db	851 CGCGTGGACCGAGTGACGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910	1870
QY		
Db	903 CCACCTCTTTGGAGGTGCGCTCTCTGCGAGCGCGCGCTTCCACCCACTCGGTGGCGCC 962	1922
QY		
Db	911 CCACCTCTTTGGAGGTGCGCTCTCTGCGAGCGCGCGCTTCCACCCACTCGGTGGCGCC 970	1930
QY		
Db	963 AGCACACGCGGGGCCCCCATCCATCGGGGACACACACCTCCCTGGGACACGCTTGTG 1022	1982
QY		
Db	971 AGCACACGCGGGGCCCCCATCCATCGGGGACACACACCTCCCTGGGACACGCTTGTG 1030	1990
QY		
Db	1023 CCCGGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGGCGCAAGAGCAGCTGC 1082	2042
QY		
Db	1031 CCCGGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCAGGCGCAAGAGCAGCTGC 1090	2050
QY		
Db	1083 GGGCCCTCTTCTACTCTCAGCTCTCTAGGCGCCAGCCTGACTGGCGCTCGAGGCTCGTGG 1142	
QY		
Db	1091 GGGCCCTCTTCTACTCTCAGCTCTCTAGGCGCCAGCCTGACTGGCGCTCGAGGCTCGTGG 1150	
QY		
Db	1143 AGACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGCACTCCCCGAGGTGGCCCCGCC 1202	
QY		
Db	1151 AGACCATCTTCTGGGTTCCAGGCCCTGGATGCCAGGCACTCCCCGAGGTGGCCCCGCC 1210	
QY		
Db	1203 TGCCCCAGCGCTACTTGGCAAAATGCGGCCCTGTTCTTGGAGCTGCTTGGGAACACACGCG 1262	
QY		
Db	1211 TGCCCCAGCGCTACTTGGCAAAATGCGGCCCTGTTCTTGGAGCTGCTTGGGAACACACGCG 1270	
QY		
Db	1263 AGTCCCTTACGGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTCAACCCAG 1322	
QY		
Db	1271 AGTCCCTTACGGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTCAACCCAG 1330	
QY		
Db	1323 CAGCGGTGTCTGTGCGGGGAGAACCCAGGCTCTGTGGCGGCCCGCCAGAGGAGG 1382	
QY		
Db	1331 CAGCGGTGTCTGTGCGGGGAGAACCCAGGCTCTGTGGCGGCCCGCCAGAGGAGG 1390	
QY		
Db	1383 ACACAGACCCCGCTGCGCTGTCAGCTGCTCCGCCAGCACAGCAGCCCTTGGCAGTGT 1442	
QY		
Db	1391 ACACAGACCCCGCTGCGCTGTCAGCTGCTCCGCCAGCACAGCAGCCCTTGGCAGTGT 1450	
QY		
Db	1443 ACAGCTTCTGCGGGGCTGCTGCGCGGCTGTGTCGCCAGGCTCTGGGGCTTCCAGGC 1502	
QY		
Db	1451 ACAGCTTCTGCGGGGCTGCTGCGCGGCTGTGTCGCCAGGCTCTGGGGCTTCCAGGC 1510	
QY		
Db	1503 ACACAGAACCGCGCTTCTCAGGAACACCAAGAAAGTTTCTTCCCTGGGAAGCATGCCA 1562	
QY		
Db	1511 ACACAGAACCGCGCTTCTCAGGAACACCAAGAAAGTTTCTTCCCTGGGAAGCATGCCA 1570	
QY		
Db	1563 AGCTCTGCTGCGAGGAGCTGACGTGGAGATGACGCTGCGGGGCTGCGCTTGGTGGCA 1622	
QY		
Db	1571 AGCTCTGCTGCGAGGAGCTGACGTGGAAATGACGCTGCGGGACTGCGCTTGGCTGGCA 1630	
QY		
Db	1623 GGAGCCAGGGGTGGCTGTGTTCCCGCCGACAGACACCGCTGCTGCGTGGAGATCCTGG 1682	
QY		
Db	1631 GGAGCCAGGGGTGGCTGTGTTCCCGCCGACAGACACCGCTGCTGCGTGGAGATCCTGG 1690	
QY		
Db	1683 CCAAGTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742	
QY		
Db	1691 CCAAGTTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750	
QY		
Db	1743 ATGTCAGGAGACACCGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGA 1802	
QY		
Db	1751 ATGTCAGGAGACACCGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGA 1810	
QY		
Db	1803 GCAAGTTGCAAGCATTGGAAATCAGACAGCACTTGAAGAGGTGACGCTCGGGAGCTGT 1862	
QY		
Db	1811 GCAAGTTGCAAGCATTGGAAATCAGACAGCACTTGAAGAGGTGACGCTCGGGAGCTGT 1870	
QY		
Db	1863 CGGAAGCAGAGGTGACGACATCGGGAGCCAGGCCCGCCCTGCTGACGCTCCAGACTCC 1922	
QY		
Db	1871 CGGAAGCAGAGGTGACGACATCGGGAGCCAGGCCCGCCCTGCTGACGCTCCAGACTCC 1930	
QY		
Db	1923 GCTTCATCCCAAGCCTGACGGCTGCGGCCGATGTTGAACATGAGTACGTCGTGGAG 1982	
QY		
Db	1931 GCTTCATCCCAAGCCTGACGGCTGCGGCCGATGTTGAACATGAGTACGTCGTGGAG 1990	
QY		
Db	1983 CCAGAACGTTTCCGAGAGAAAGAGGCGCGCTCTCAGCTCGAGGTGAAGCACTGT 2042	
QY		
Db	1991 CCAGAACGTTTCCGAGAGAAAGAGGCGCGCTCTCAGCTCGAGGTGAAGCACTGT 2050	
QY		

Db 2043 TCAGCGTCTCAACTACGAGCGGGCGCGCCCGCCCTCTCTGGCGCCTCTGTGCTGG 2102  
Qy 2051 TCAGCGTCTCAACTACGAGCGGGCGCGCCCGCCCTCTCTGGCGCCTCTGTGCTGG 2110  
Db 2103 GCCTGGAGCATATCCACAGGGCCTGGCGACCTTCGTGCTGCGTGGCGGCCAGGACC 2162  
Qy 2111 GCCTGGAGCATATCCACAGGGCCTGGCGACCTTCGTGCTGCGTGGCGGCCAGGACC 2170  
Db 2163 GCCTGGAGCATATCTTCAAGTGGATGTGACGGGGCGGTACGACACCATCCCCC 2222  
Qy 2171 GCCTGGAGCATATCTTCAAGTGGATGTGACGGGGCGGTACGACACCATCCCCC 2230  
Db 2223 AGGACAGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGACAGCTACTGCGTGC 2282  
Qy 2231 AGGACAGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGACAGCTACTGCGTGC 2290  
Db 2283 GTGCGTATGCGGTGTCAGAGGGCGCCCATGCGGACGTGCGCAAGCCCTCAAGAGCC 2342  
Qy 2291 GTGCGTATGCGGTGTCAGAGGGCGCCCATGCGGACGTGCGCAAGCCCTCAAGAGCC 2350  
Db 2343 AGCTCTTACCTTGACAGACCTCCAGCCGTATCATCGAGCAGCTCTCCCTGAATGAGGCA 2402  
Qy 2351 AGCTCTTACCTTGACAGACCTCCAGCCGTATCATCGAGCAGCTCTCCCTGAATGAGGCA 2410  
Db 2403 AGACAGCGCCGTGAGGATGCGGTGTCATCGAGCAGCTCTCCCTGAATGAGGCA 2462  
Qy 2411 AGACAGCGCCGTGAGGATGCGGTGTCATCGAGCAGCTCTCCCTGAATGAGGCA 2470  
Db 2463 GCAGTGGCCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCT 2522  
Qy 2471 GCAGTGGCCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCT 2530  
Db 2523 GCAGTGGCCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCT 2582  
Qy 2531 GCAGTGGCCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCTTCCAGCTCT 2590  
Db 2593 GCAGCTGTGTACGGCAGCATGAGACAGAGCTGTTGGGGGATTCGGGGGACGGGC 2642  
Qy 2591 GCAGCTGTGTACGGCAGCATGAGACAGAGCTGTTGGGGGATTCGGGGGACGGGC 2650  
Db 2643 TGCTCTCGTTTGTGATGATTCTTGTGTCACACCTCACCTCACCCACCGCAAAA 2702  
Qy 2651 TGCTCTCGTTTGTGATGATTCTTGTGTCACACCTCACCTCACCCACCGCAAAA 2710  
Db 2703 CTTCTCTCAGGACCTGTCAGAGTGTCCCTGAGTATGGCTGCTGCTGCTGCTGCTGCTG 2762  
Qy 2711 CTTCTCTCAGGACCTGTCAGAGTGTCCCTGAGTATGGCTGCTGCTGCTGCTGCTGCTG 2770  
Db 2763 AGACAGTGTGNACTTCCCTGTAGAGAGAGGCGCCCTGGGTGGCAGCGCTTTGTTTCA 2822  
Qy 2771 AGACAGTGTGNACTTCCCTGTAGAGAGAGGCGCCCTGGGTGGCAGCGCTTTGTTTCA 2830  
Db 2823 TGCGGGCCACGGCCTAT 2840  
Qy 2831 TGCGGGCCACGGCCTAT 2848

RESULT 8  
ID US-08-854-050-173 STANDARD; DNA; UNC; 4029 BP.  
AC xxxxxx  
DT

Sequence 173, Application US/08854050  
Sequence 173, Application US/08854050  
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Novel Telomerase  
NUMBER OF SEQUENCES: 225

CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 173:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4029 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY:  
CC LOCATION: 1..4029  
CC OTHER INFORMATION: /note= "preliminary sequence for  
CC OTHER INFORMATION: human TRT cDNA insert of  
CC OTHER INFORMATION: plasmid pGRN121"  
CC SEQUENCE 4029 BP; 687 A; 1342 C; 1255 G; 720 T; 25 OTHER.

Query Match 88.5%; Score 2521; DB 21; Length 4029;  
Best Local Similarity 97.0%; Pred. No. 0.00e+00;  
Matches 2754; Conservative 0; Mismatches 77; Indels 7; Gaps 6;

Db 1 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 60  
Qy 12 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 71  
Db 61 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 120  
Qy 72 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 131  
Db 121 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 180  
Qy 132 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 191  
Db 181 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 240  
Qy 192 GCAGCGCTGGCTCTGCTGGCGACGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 251







Db	659	CTGGAACCATAGCGTACAGGAGCGCGGGTCCCTCTGGCC-TGCCAGCCCGGGTGGCG	717
Qy	672	CTGGAACCATAGCGTACAGGAGCGCGGGTCCCTCTGGCC-TGCCAGCCCGGGTGGCG	731
Db	718	GAGCGCGGGGCGAGTGCACAGCCGAAGTCTGCGGTGCGCCAAAGAGGCCAGGCGTGGCG	777
Qy	732	GAGCGCGGGGCGAGTGCACAGCCGAAGTCTGCGGTGCGCCAAAGAGGCCAGGCGTGGCG	791
Db	778	TGCCCCTAGCGCGGAGCGGACCCCGTTGGCGAGGGGTCTTGGGCCACACCCGGCGAGGAC	837
Qy	792	TGCCCCTAGCGCGGAGCGGACCCCGTTGGCGAGGGGTCTTGGGCCACACCCGGCGAGGAC	851
Db	838	GCCTGACCGAGTGACCGTGTCTGTGTGTGTACTTGCCTGCCAGACCCGCCGAAGAAGC	897
Qy	852	GCCTGACCGAGTGACCGTGTCTGTGTGTGTACTTGCCTGCCAGACCCGCCGAAGAAGC	911
Db	898	CACCTCTTTGGAGGFGCGCTCTTGGCAGCGGCACATCCACCACATCCGTGGCGCGCCA	957
Qy	912	CACCTCTTTGGAGGFGCGCTCTTGGCAGCGGCACATCCACCACATCCGTGGCGCGCCA	971
Db	958	GCACCACGGGGGCCCCCATCCACATCGCGGCCACACAGTCC-TGGGACACAGCCTTGTCC	1016
Qy	972	GCACCACGGGGGCCCCCATCCACATCGCGGCCACACAGTCC-TGGGACACAGCCTTGTCC	1031
Db	1017	CCCGGTGTAGCGCGAGACCAAGCACTTCTCTACTTCTCTAGGCGACAGA-ACTCTCGN	1075
Qy	1032	CCCGGTGTAGCGCGAGACCAAGCACTTCTCTACTTCTCTAGGCGACAGAAGCAGCTGCG	1091
Db	1076	CCCTCCTTCTACTCAATATATCTGAGGCCACAGCTGACTGGCGTTGGGAGGTTCTGTGG	1135
Qy	1092	GCCTCCTTCTACTGAGCTCTTCTGAGGCCACAGCTGACTGGCGTTGGGAGGTTCTGTGG	1151
Db	1136	AGACANTCTTTCTGGTTCCAGGCGCTTGGATGCCAGG-ATTCCC CGCAGGTTGCCCGGCT	1194
Qy	1152	GACCATCTTTCTGGGTTCCAGGCGCTTGGATGCCAGGAGTCTCCCGCAGGTTGCCCGGCT	1211
Db	1195	GCCCCAGGNTACTTGCAAAATCGGCGCCCTTGTCTGAGGTGTGTGGGAACACAGCGCA	1254
Qy	1212	GCCCCAGGCGTACTTGGCAAAATCGGCGCCCTGTTTCTGAGGTGTGTGGGAACACAGCGCA	1271
Db	1255	GTGCCCTTACGGGGTGTCTCTCAAGACGACTGCCCGCTCGGAGCTGGGCTCACCCGAGC	1314
Qy	1272	GTGCCCTTACGGGGTGTCTCTCAAGACGACTGCCCGCTCGGAGCTGGGCTCACCCGAGC	1331
Db	1315	AGCGGTGTCTGTGCCGGGAGAACCCAGGGCTCTGTGCGGCCCGCCAGAGGAGGAGA	1374
Qy	1332	AGCGGTGTCTGTGCCGGGAGAACCCAGGGCTCTGTGCGGCCCGCCAGAGGAGGAGA	1391
Db	1375	ACACAGACCCCGTGGCTGTGTGTCAGCTGTCTCGCCAGCAGACAGACCCCTTGCGAGTGT	1434
Qy	1392	-CACAGACCCCGTGGCTGTGTGTCAGCTGTCTCGCCAGCAGACAGACCCCTTGCGAGTGT	1450
Db	1435	ACGGCTTCTGTGGGGCTTGCCTTGCAGGAGTGTGTCAGGCTGTGGGGCTTCAGGC	1494
Qy	1451	ACGGCTTCTGTGGGGCTTGCCTTGCAGGAGTGTGTCAGGCTGTGGGGCTTCAGGC	1510
Db	1495	ACAACGAACCGCGCTTCTTCAGGAACACCAAGAAGTTTCATCTCCCTGGGAAGCATGCCA	1554
Qy	1511	ACAACGAACCGCGCTTCTTCAGGAACACCAAGAAGTTTCATCTCCCTGGGAAGCATGCCA	1570
Db	1555	AGCTCTCGTGCAGGAGCTGACGTGGAAGTGTGCGGAGTGTGCGGACTGCGCTTGGTGGCGCA	1614
Qy	1571	AGCTCTCGTGCAGGAGCTGACGTGGAAGTGTGCGGAGTGTGCGGACTGCGCTTGGTGGCGCA	1630
Db	1615	GGAGCCAGGGGTTGGCTGTGTTCGGCGCGCAGACACCGTCTCGGTGAGGAGATCTCTGG	1674
Qy	1631	GGAGCCAGGGGTTGGCTGTGTTCGGCGCGCAGACACCGTCTCGGTGAGGAGATCTCTGG	1690
Db	1675	CCAAGTTCTGTGACTGGCTGATGAGTGTGTAGTGTGTAGTGTGTAGTGTGTCTCTCTTT	1734
Qy	1691	CCAAGTTCTGTGACTGGCTGATGAGTGTGTAGTGTGTAGTGTGTAGTGTGTCTCTCTTT	1750
Db	1735	ATGTCAGGAGACACCGTTTCAAAAGAACAGAGCTCTTTTTTCTACCGGAAGAGTGTCTGGA	1794

QY	1751	ATGTCACGGAGACCACCGTCTTTCAAAGAAGACGGCTCTTTTTTCTACCCGAAGAGTGCTGTGGA	1810
Db	1795	GCAAGTTGCAAAAGCATTGGAATCAGACACACTTGAAGAGGGTGACGTGCGGAGAGCTGT	1854
QY	1811	GCAAGTTGCAAAAGCATTGGAATCAGACACACTTGAAGAGGGTGACGTGCGGAGAGCTGT	1870
Db	1855	CGAAGCAGAGGTCAGGCAGCATCGGAAGACCGGCCCGCTGTGACGTCCAGACTCC	1914
QY	1871	CGAAGCAGAGGTCAGGCAGCATCGGAAGACCGGCCCGCTGTGACGTCCAGACTCC	1930
Db	1915	GCTTCATCCCAGCCCTGACGGGCTCGGGCGGATGTGAACTATGACTACGTGCTGGGAG	1974
QY	1931	GCTTCATCCCAGCCCTGACGGGCTCGGGCGGATGTGAACTATGACTACGTGCTGGGAG	1990
Db	1975	CCAGAAGCTTCCGAGAGAAAAGAGGGCCGACGCTCTCACCTCGAGSGGTGAAGCAGCTGT	2034
QY	1991	CCAGAAGCTTCCGAGAGAAAAGAGGGCCGACGCTCTCACCTCGAGSGGTGAAGCAGCTGT	2050
Db	2035	TCAGCGTGCTCAACTAGAGCGGGCGGGCCCGGGCTCTCGGGCGGCTCTGTGCTGG	2094
QY	2051	TCAGCGTGCTCAACTAGAGCGGGCGGGCCCGGGCTCTCGGGCGGCTCTGTGCTGG	2110
Db	2095	GCCTGGAGGATATCCACAGGGCTGCGGCACCTTCGTGCTGGTGTGCGGGCCCCAGGACC	2154
QY	2111	GCCTGGAGGATATCCACAGGGCTGCGGCACCTTCGTGCTGGTGTGCGGGCCCCAGGACC	2170
Db	2155	CGCCGCTGAGCTGTACTTTGTCAAAGTGGAATGTGACGGGGCGGTACGACACCATCCCCC	2214
QY	2171	CGCCGCTGAGCTGTACTTTGTCAAAGTGGAATGTGACGGGGCGGTACGACACCATCCCCC	2230
Db	2215	AGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCAGAACAGCTACTGCGTGC	2274
QY	2231	AGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCAGAACAGCTACTGCGTGC	2290
Db	2275	GTGCGTATGCGTGGTCCAGAAGGCCGCCCATGGGACGTTCGCGAAGGCCCTTCAAGAGCC	2334
QY	2291	GTGCGTATGCGTGGTCCAGAAGGCCGCCCATGGGACGTTCGCGAAGGCCCTTCAAGAGCC	2350
Db	2335	ACGTCTCTACTTGTACAGACCTCCAGCCGTACATCGGACAGTTCTGTGCGTCACCTGCAGG	2394
QY	2351	ACGTCTCTACTTGTACAGACCTCCAGCCGTACATCGGACAGTTCTGTGCGTCACCTGCAGG	2410
Db	2395	ANACAGCCCGCTGAGGGATGCGGTGCTCATCGAGCAGAGCTCCCTCGATGAAGGCCA	2454
QY	2411	AGACAGCCCGCTGAGGGATGCGGTGCTCATCGAGCAGAGCTCCCTCGATGAAGGCCA	2470
Db	2455	CGAGTGGCGCTTCGAGCGTCTTCTACGCTTCATGTCACACACCGCTGCGCATCAGGG	2514
QY	2471	CGAGTGGCGCTTCGAGCGTCTTCTACGCTTCATGTCACACACCGCTGCGCATCAGGG	2530
Db	2515	GCNAGTCTTAGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCT	2574
QY	2531	GCNAGTCTTAGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCT	2590
Db	2575	GCAGCCTGTGTACGGCGACATGGAGAAACAAGCTGTTGCGGGGATTCGGCGGGACGGC	2634
QY	2591	GCAGCCTGTGTACGGCGACATGGAGAAACAAGCTGTTGCGGGGATTCGGCGGGACGGC	2650
Db	2635	TGCTCTCGGTTTGGTGGATGATTTCTTGTGTTGACACCTCACCTCACCCACCGGAAAA	2694
QY	2651	TGCTCTCGGTTTGGTGGATGATTTCTTGTGTTGTTGACACCTCACCTCACCCACCGGAAAA	2710
Db	2695	CCTTCTCAGAACCCCTGGTCCGAGGTGTCCTGTAGTATGGCTGGTGGTGAACCTTCGCGA	2754
QY	2711	CCTTCTCAGAACCCCTGGTCCGAGGTGTCCTGTAGTATGGCTGGTGGTGAACCTTCGCGA	2770
Db	2755	AGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCGACGCTTTGTTTACA	2814
QY	2771	AGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCGACGCTTTGTTTACA	2830
Db	2815	TGCGGGGCCACGGCCTAT	2832





QY 972 GACACAGGGGGCCCCCATCCACATCGGGCCACACAGTCCCTGGGACACGCCCTGTGCC 1031  
Db 1021 CCCGGTGACGCCGAGAGCAAGCACTTCTCTACTCTCAGCGCAACAAGGACAGCTCGG 1080  
QY 1032 CCCGGTGACGCCGAGAGCAAGCACTTCTCTACTCTCAGCGCAACAAGGACAGCTCGG 1091  
Db 1081 GCGCTCTCTCTACTCTCTCTCAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGA 1140  
QY 1092 GCGCTCTCTCTACTCTCTCTCAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGA 1151  
Db 1141 GACCATCTTCTGGGTTCAGGCCCTGGATGCCAGGGACTCCCGCAGGTTCGCCCGGCT 1200  
QY 1152 GACCATCTTCTGGGTTCAGGCCCTGGATGCCAGGGACTCCCGCAGGTTCGCCCGGCT 1211  
Db 1201 GCGCCAGCGCTACTGGCAAAATGCGCCCTGTTTCTGAGCTGCTTGGGAACACGCGCA 1260  
QY 1212 GCGCCAGCGCTACTGGCAAAATGCGCCCTGTTTCTGAGCTGCTTGGGAACACGCGCA 1271  
Db 1261 GTGCCCCCTACGGGGTGCTCTCAAGACGACTGCCCGCTGCGAGCTGGGTCAACCCAGC 1320  
QY 1272 GTGCCCCCTACGGGGTGCTCTCAAGACGACTGCCCGCTGCGAGCTGGGTCAACCCAGC 1331  
Db 1321 AGCCGGTGCTGTCGCCGGGAAAGCCCGAGGGCTCTGTGGCGGCCCGGAGGAGGA 1380  
QY 1332 AGCCGGTGCTGTCGCCGGGAAAGCCCGAGGGCTCTGTGGCGGCCCGGAGGAGGA 1391  
Db 1381 CACAGACCCCCCTCGCTGGTGCAGTCTCTCGCCAGCACAGCAGCCCTGGCAGGTGA 1440  
QY 1392 CACAGACCCCCCTCGCTGGTGCAGTCTCTCGCCAGCACAGCAGCCCTGGCAGGTGA 1451  
Db 1441 CGGGTTCTGTCGGGGCTGCTCGCGGGTGTGTGCCCGCCAGGCCCTCTGGGGTTCAGGCA 1500  
QY 1452 CGGGTTCTGTCGGGGCTGCTCGCGGGTGTGTGCCCGCCAGGCCCTCTGGGGTTCAGGCA 1511  
Db 1501 CAACGAAGCCGCTTCTCAGAACACCAAGAGTTCATCTCCCTGGGAGCATGCCAA 1560  
QY 1512 CAACGAAGCCGCTTCTCAGAACACCAAGAGTTCATCTCCCTGGGAGCATGCCAA 1571  
Db 1561 GCTCTCGTGCAGGAGTGCAGTGAAGATGACGCTGGGGACTGCGCTTGGCTGGCAG 1620  
QY 1572 GCTCTCGTGCAGGAGTGCAGTGAAGATGACGCTGGGGACTGCGCTTGGCTGGCAG 1631  
Db 1621 GAGCCAGGGGTGGTGTGTTCCGGCCGACAGCACGCTCTGCGTGAAGAGATCCTGCG 1680  
QY 1632 GAGCCAGGGGTGGTGTGTTCCGGCCGACAGCACGCTCTGCGTGAAGAGATCCTGCG 1691  
Db 1681 CAAGTCTCTGCACTGGCTGATGAGTGTGACGCTCGAGCTGCTCAGTCTTCTTTTA 1740  
QY 1692 CAAGTCTCTGCACTGGCTGATGAGTGTGACGCTCGAGCTGCTCAGTCTTCTTTTA 1751  
Db 1741 TGTACGGAGACACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1800  
QY 1752 TGTACGGAGACACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1811  
Db 1801 CAAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGGGGTGACGCTCGGGAGCTGTC 1860  
QY 1812 CAAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGGGGTGACGCTCGGGAGCTGTC 1871  
Db 1861 GGAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCCCGCTGCTGACGCTCCAGACTCGG 1920  
QY 1872 GGAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCCCGCTGCTGACGCTCCAGACTCGG 1931  
Db 1921 CTTTCATCCCAAGCCTGACGGCTGGGGCCGATTGTGAACATGGACTAGCTCGTGGGAGC 1980  
QY 1932 CTTTCATCCCAAGCCTGACGGCTGGGGCCGATTGTGAACATGGACTAGCTCGTGGGAGC 1991  
Db 1981 CAGAAGCTTCCGAGAGAAAGAGGCCGAGGCTCTCACCCTCGAGGGTGAAGGCACTGTT 2040  
QY 1992 CAGAAGCTTCCGAGAGAAAGAGGCCGAGGCTCTCACCCTCGAGGGTGAAGGCACTGTT 2051  
Db 2041 CAGCGTCTCACTAGCAGCGGGGGGGCGGCCCTCTCTGGGGCGCTCTGTGCTGGG 2100

QY 2052 CAGCGTCTCACTACGAGCGGGCGCGGCCCTCTCTGGGGCGCTCTGTGCTGGG 2111  
Db 2101 CCTGGACGATATCCACAGGGCTGGCGACCTTCTGTCTGCTGTGCGGGCCCCAGGACCC 2160  
QY 2112 CTTGACGATATCCACAGGGCTGGCGACCTTCTGTCTGCTGTGCGGGCCCCAGGACCC 2171  
Db 2161 GCGCGCTCAGCTGACTTGTCAAGGTGGATGTGACGGGGCGGTACGACACCATCCGCCA 2220  
QY 2172 GCGCGCTCAGCTGACTTGTCAAGGTGGATGTGACGGGGCGGTACGACACCATCCGCCA 2231  
Db 2221 GGACAGGCTCAGGAGTTCATCGCAGCATCATCAAAACCCAGAACACGTACTTGGTGGC 2280  
QY 2232 GGACAGGCTCAGGAGTTCATCGCAGCATCATCAAAACCCAGAACACGTACTTGGTGGC 2291  
Db 2281 TCGGTATGCCGTGTGCCAAGAGCGGCCCATGGGACGTCGCCAAGGCTTCAAGAGCCA 2340  
QY 2292 TCGGTATGCCGTGTGCCAAGAGCGGCCCATGGGACGTCGCCAAGGCTTCAAGAGCCA 2351  
Db 2341 CGTC----- 2344  
QY 2352 CGTCTTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGTCACTGTCAGGA 2411  
Db 2344 ----- 2344  
QY 2412 GACAGCGCGCTGAGGGATGCCGTGCTCATCGAGCAGAGCTCTCCCTGAAATGAGGCCAG 2471  
Db 2344 ----- 2344  
QY 2472 CAGTGGCGCTTTCGACGCTTCTTACGCTTCATGTGCCACACGCGCTGGCATCAGGGG 2531  
Db 2345 -----CTACGCTCAGTCCAGGGGATCCCGAGGGTTCATTCCTCCACGCTGCTGTG 2398  
QY 2532 CAACTCTTACGCTCAGTCCAGGGGATCCCGAGGGTTCATTCCTCCACGCTGCTGTG 2591  
Db 2399 CAGCGTGTCTACGCGCATGGAGAACAAAGCTGTTTCGGGGATTCGCGGGAGCGGCT 2458  
QY 2592 CAGCGTGTCTACGCGCATGGAGAACAAAGCTGTTTCGGGGATTCGCGGGAGCGGCT 2651  
Db 2459 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGCGAAAC 2518  
QY 2652 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGCGAAAC 2711  
Db 2519 CTTCTCAGGACCTGTGTCGAGGTGTCCTTGAGTATGGTGTGCTGTGAGTTCGCGAA 2578  
QY 2712 CTTCTCAGGACCTGTGTCGAGGTGTCCTTGAGTATGGTGTGCTGTGAGTTCGCGAA 2771  
Db 2579 GACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGTGGCAGCGGCTTTGTTTCAGAT 2638  
QY 2772 GACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGTGGCAGCGGCTTTGTTTCAGAT 2831  
Db 2639 GCGGCGCCACGCGCTAT 2655  
QY 2832 GCGGCGCCACGCGCTAT 2848

## RESULT 11

ID US-08-912-951-4 STANDARD; DNA; UNC; 3855 BP.

AC xxxxxx

DT

DE Sequence 4, Application US/08912951

CC Sequence 4, Application US/08912951

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC NUMBER OF SEQUENCES: 335

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC City: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 4:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3855 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: -  
CC LOCATION: 1..3855  
CC OTHER INFORMATION: /note= "nucleic acid sequence with an  
CC OTHER INFORMATION: open reading frame encoding a delta-182  
CC OTHER INFORMATION: variant polypeptide"  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56..2479  
CC OTHER INFORMATION: /product= "delta-182 variant  
CC OTHER INFORMATION: polypeptide"  
CC OTHER INFORMATION: polypeptide"  
CC SEQUENCE 3855 BP; 651 A; 1300 C; 1226 G; 678 T; 0 OTHER.

Query Match 85.8%; Score 2444; DB 22; Length 3855;  
Best Local Similarity 93.6%; Pred. No. 0.00e+00;  
Matches 2655; Conservative 0; Mismatches 0; Indels 182; Gaps 1;  
DB 1 GCAGCGCTGGCTCTGCTGGGACAGTGGGAAGCCCTGGGCCCCGGGCCCCGCGATGCC 60  
QY 12 GCAGCGCTGGCTCTGCTGGGACAGTGGGAAGCCCTGGGCCCCGGGCCCCGCGATGCC 71  
DB 61 GGGCGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGGCGCAGCAGTACCAGGAGTGCT 120  
QY 72 GCGCGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGGCGCAGCAGTACCAGGAGTGCT 131

DB 121 GCCGCTGCCACAGTTCGTGCGGCGCTTGGGCGCCCAAGGGCTGGGCGCTGGTGACGCGGG 180  
QY 132 GCCGCTGCCACAGTTCGTGCGGCGCTTGGGCGCCCAAGGGCTGGGCGCTGGTGACGCGGG 191  
DB 181 GGACCCGGCGCTTTCCGCGCGCTGGTGCGCCAGTCCGCTGCTGCTGCTGGGAGCG 240  
QY 192 GGACCCGGCGCTTTCCGCGCGCTGGTGCGCCAGTCCGCTGCTGCTGCTGGGAGCG 251  
DB 241 ACGGCGCGCCCGCGCGCGCTTCTTCCGCGAGTGTCTCTGCTGAAAGAGCTGGTGCG 300  
QY 252 ACGGCGCGCCCGCGCGCGCTTCTTCCGCGAGTGTCTCTGCTGAAAGAGCTGGTGCG 311  
DB 301 CCGAGTGTCTGAGAGCTGTGCGAGCGCGCGCGGGAAGAACGTGCTGCGCTTCGCTGCG 360  
QY 312 CCGAGTGTCTGAGAGCTGTGCGAGCGCGCGCGGGAAGAACGTGCTGCGCTTCGCTGCG 371  
DB 361 GCTGTGACGCGCGCGCGCGCGCTTCCGCGAGTGTCTCTGCTGAAAGAGCTGGTGCG 420  
QY 372 GCTGTGACGCGCGCGCGCGCGCTTCCGCGAGTGTCTCTGCTGAAAGAGCTGGTGCG 431  
DB 421 CCTGCCAACACAGTGTACCGACGCTGCGGGGAGCGGGGCTGCTGCTGCG 480  
QY 432 CCTGCCAACACAGTGTACCGACGCTGCGGGGAGCGGGGCTGCTGCTGCG 491  
DB 481 CCGGCTGGGCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 492 CCGGCTGGGCGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551  
DB 541 GGTCTCCAGCTGGCGCTACAGTGTGCGGCGCGCGCTGTACAGCTGCGGCTGCGCAC 600  
QY 552 GGTCTCCAGCTGGCGCTACAGTGTGCGGCGCGCGCTGTACAGCTGCGGCTGCGCAC 611  
DB 601 TCAGCGCGCGCGCGCGCTAGTGTGAGCCCGGAGGCGTCTGGATGCGAACGGCG 660  
QY 612 TCAGCGCGCGCGCGCGCTAGTGTGAGCCCGGAGGCGTCTGGATGCGAACGGCG 671  
DB 661 CTGGAACCATAGCGTACGAGGCGCGGGTCCCGCTTGGGCTGCGAGCCCGGGTGGCG 720  
QY 672 CTGGAACCATAGCGTACGAGGCGCGGGTCCCGCTTGGGCTGCGAGCCCGGGTGGCG 731  
DB 721 GAGCGCGGGGCGAGTGCACCGAAGTCTGCGCTTGGGCTGCGAGCCCGGGTGGCG 780  
QY 732 GAGCGCGGGGCGAGTGCACCGAAGTCTGCGCTTGGGCTGCGAGCCCGGGTGGCG 791  
DB 781 TGCCCTGAGCGGAGCGAGCGCGCTTGGGCTGCGAGCCCGGGTGGCG 840  
QY 792 TGCCCTGAGCGGAGCGAGCGCGCTTGGGCTGCGAGCCCGGGTGGCG 851  
DB 841 GCGTGGACCGAGTGACCGTGGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 852 GCGTGGACCGAGTGACCGTGGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911  
DB 901 CACTCTTTGAGGAGTGGCTCTGCGACGCGCGCTTCCGCGAGTCCGCTGCGCGCGCA 960  
QY 912 CACTCTTTGAGGAGTGGCTCTGCGACGCGCGCTTCCGCGAGTCCGCTGCGCGCGCA 971  
DB 961 GCACACGCGCGCGCGCGCTTCCGCGAGTCCGCGAGTCCGCGAGTCCGCGAGTCCGCGAGT 1020  
QY 972 GCACACGCGCGCGCGCGCTTCCGCGAGTCCGCGAGTCCGCGAGTCCGCGAGTCCGCGAGT 1031  
DB 1021 CCGGCTGTACGCGGAGCAAGACATTCCTTCTACTCTCTAGGCGGCAAGAGGAGTCCG 1080  
QY 1032 CCGGCTGTACGCGGAGCAAGACATTCCTTCTACTCTCTAGGCGGCAAGAGGAGTCCG 1091  
DB 1081 GCCCTCTTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1140  
QY 1092 GCCCTCTTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1151  
DB 1141 GACCATCTTCTGGGTTCCAGGCGCTGATGCGAGGAGTCCCGCAGGAGTCCCGCGCT 1200  
QY 1152 GACCATCTTCTGGGTTCCAGGCGCTGATGCGAGGAGTCCCGCAGGAGTCCCGCGCT 1211

Db 1201 GCGCCAGGCTACTGGCAATCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA 1260  
QY 1212 GCGCCAGGCTACTGGCAATCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCA 1271  
Db 1261 GTGCCCTACGGGTGCTCTCAAGAGCGACTGCCCTCGAGCTCGGTTCACCCAGC 1320  
QY 1272 GTGCCCTACGGGTGCTCTCAAGAGCGACTGCCCTCGAGCTCGGTTCACCCAGC 1331  
Db 1321 AGCGGTGTGTGCGCGGAGAGCCCAAGGGCTCTGTGGCGGCCCCCGAGGAGGA 1380  
QY 1332 AGCGGTGTGTGCGCGGAGAGCCCAAGGGCTCTGTGGCGGCCCCCGAGGAGGA 1391  
Db 1381 CACAGACCCCGCTGCTGTGCAAGTCTCCGCCACACAGACGCCCTGGCAGTGT 1440  
QY 1392 CACAGACCCCGCTGCTGTGCAAGTCTCCGCCACACAGACGCCCTGGCAGTGT 1451  
Db 1441 CGGCTTGTGCGGCCCTGCTGCGCGGCTGGTGCCGCCAGGCTCTGGGGCTCCAGCA 1500  
QY 1452 CGGCTTGTGCGGCCCTGCTGCGCGGCTGGTGCCGCCAGGCTCTGGGGCTCCAGCA 1511  
Db 1501 CAAGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCCCTGGGGAAGCATGCCAA 1560  
QY 1512 CAAGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCCCTGGGGAAGCATGCCAA 1571  
Db 1561 GCTCTCCTGCAGAGCTGAGCTGGAAGATCAGCTGCGGGACTCGCTTGGCTGCCAG 1620  
QY 1572 GCTCTCCTGCAGAGCTGAGCTGGAAGATCAGCTGCGGGACTCGCTTGGCTGCCAG 1631  
Db 1621 GAGCCAGGGTGGCTGTGTTCGGCGCGAGACACCGCTCTGGTGAGGAGATCTCTGC 1680  
QY 1632 GAGCCAGGGTGGCTGTGTTCGGCGCGAGACACCGCTCTGGTGAGGAGATCTCTGC 1691  
Db 1681 CAAGTTCTCCTGCTGCTGATGAGTGTGTACGTCTGCTGAGTGTCTCAGGTCTTTCTTTA 1740  
QY 1692 CAAGTTCTCCTGCTGCTGATGAGTGTGTACGTCTGCTGAGTGTCTCAGGTCTTTCTTTA 1751  
Db 1741 TGTACAGGACACAGCTTTCAAAGACAGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1800  
QY 1752 TGTACAGGACACAGCTTTCAAAGACAGCTCTTTTCTACCGGAAGAGTGTCTGGAG 1811  
Db 1801 CAAGTTCAAAGCATTTGGAATCAGACAGCACTTGAAGAGGTGCTGCGGGAGCTGTC 1860  
QY 1812 CAAGTTCAAAGCATTTGGAATCAGACAGCACTTGAAGAGGTGCTGCGGGAGCTGTC 1871  
Db 1861 GGAAGCAGAGTCAAGCAGATCGGGAAGCCAGCGCCCTGCTGACGTCACAGCTCCG 1920  
QY 1872 GGAAGCAGAGTCAAGCAGATCGGGAAGCCAGCGCCCTGCTGACGTCACAGCTCCG 1931  
Db 1921 CTTTCATCCCAAGCTGACGGGCTGCGGCGATTTGTGAACATGACACGTCGTTGGAGC 1980  
QY 1932 CTTTCATCCCAAGCTGACGGGCTGCGGCGATTTGTGAACATGACACGTCGTTGGAGC 1991  
Db 1981 CAGAACGTTCCGAGAGAAAGAGGCGCGCGCTCTCCTGCGCGCTCTGTCGTGGG 2040  
QY 1992 CAGAACGTTCCGAGAGAAAGAGGCGCGCGCTCTCCTGCGCGCTCTGTCGTGGG 2051  
Db 2041 CAGCGTGTCTCAACTACAGCGGCGCGCGCGCTCTCCTGCGCGCTCTGTCGTGGG 2100  
QY 2052 CAGCGTGTCTCAACTACAGCGGCGCGCGCGCTCTCCTGCGCGCTCTGTCGTGGG 2111  
Db 2101 CTTGAGCATATCACAAGGCTGCGGCGACTTCTGCTGCTGCGGCGCCAGGACCC 2160  
QY 2112 CTTGAGCATATCACAAGGCTGCGGCGACTTCTGCTGCTGCGGCGCCAGGACCC 2171  
Db 2161 GCGGCTGAGCTGACTTGTCAAGTGGATGTACGGGCGCGTACGACACCATCCGCCA 2220  
QY 2172 GCGGCTGAGCTGACTTGTCAAGTGGATGTACGGGCGCGTACGACACCATCCGCCA 2231  
Db 2221 GGACAGGCTCAGGAGGTCTATCGCAGCATCATCAAAACCCAGAACACGACTCGTGGC 2280  
QY 2232 GGACAGGCTCAGGAGGTCTATCGCAGCATCATCAAAACCCAGAACACGACTCGTGGC 2291  
Db 2281 TCGGTATGCGGTGTCTCCAGAGGCGGCCCATGGGCGAGCTCGGCAAGGCTTCAAGAGCCA 2340

QY 2292 TCAGTATGCGGTGGTCCAGAAAGCCGCCATGGCGACGTCGCCAAGGCCTTCAAGAGCCA 2351  
Db 2341 CGTC-----CTAGCTCAGTGGCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2344  
QY 2352 CGTCTACCTTGACAGACCTCCAGCCGACATGCGACAGTTGCTGGCTCACCTGCAGGA 2411  
Db 2344 -----GCTCCTAGTCCAGTCCGAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2344  
QY 2412 GACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCTGAATGAGGCCAG 2471  
Db 2344 -----CAGTGGCTCTTCGAGCTCTTCCTACGTTCTATGTGCGCACCAGCCGTCGCGATCAGGGG 2344  
QY 2472 CAGTGGCTCTTCGAGCTCTTCCTACGTTCTATGTGCGCACCAGCCGTCGCGATCAGGGG 2531  
Db 2345 -----CTAGCTCAGTGGCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2398  
QY 2532 CAAGTCTCAGTCCAGTCCGAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2591  
Db 2399 CAGCCTGTGTACGGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2458  
QY 2592 CAGCCTGTGTACGGGACATGGAGAACAGCTGTTTGGGGGATTCGGGGGACGGGCT 2651  
Db 2459 GCTCCTGCGTTCGATGATTTCTTGTGTGACACCTCACCCTCACCCACCGGAAAC 2518  
QY 2652 GCTCCTGCGTTCGATGATTTCTTGTGTGACACCTCACCCTCACCCACCGGAAAC 2711  
Db 2519 CTTCTCAGGACCCCTGCTCGAGGTGTCCTGAGTATGGCTGGTGGTGAACCTTGGGAA 2578  
QY 2712 CTTCTCAGGACCCCTGCTCGAGGTGTCCTGAGTATGGCTGGTGGTGAACCTTGGGAA 2771  
Db 2579 GACAGTGGTGAACCTCCCTGTAGAAGCAGGCGCTGGTGGCAGCGCTTTTGTTCAGAT 2638  
QY 2772 GACAGTGGTGAACCTCCCTGTAGAAGCAGGCGCTGGTGGCAGCGCTTTTGTTCAGAT 2831  
Db 2639 GCCGGCCCGGCGCTAT 2655  
QY 2832 GCCGGCCCGGCGCTAT 2848

RESULT 12  
ID US-08-912-951-6 STANDARD; DNA; UNC; 4200 BP.  
AC xxxxxx  
DT Sequence 6, Application US/08912951  
CC Sequence 6, Application US/08912951  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC NUMBER OF SEQUENCES: 335 THERAPEUTIC METHODS  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435



QY 1467 CTGCTGCGCGGCTGTGCTCCCGCAGGCTCTGGGGCTCCAGGCACACGACGCCGCTT 1526  
Db 3877 CCTAGAACACCAAGATTCTCCTCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGA 3936  
QY 1527 CCTAGGAACACCAAGATTCTCCTCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGA 1586  
Db 3937 GCTGACGTGGAGATGAGCGTGGGACTGCGCTTGGCTGGCGAGGCCAGG 3990  
QY 1587 GCTGACGTGGAGATGAGCGTGGGACTGCGCTTGGCTGGCGAGGCCAGG 1640

RESULT 13  
ID US-09-042-460-1 STANDARD; DNA; UNC; 3496 BP.  
AC xxxxxx

DE Sequence 1, Application US/09042460  
CC Sequence 1, Application US/09042460  
CC GENERAL INFORMATION:  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Allsopp, Richard  
CC APPLICANT: DePinho, Ronald  
CC APPLICANT: Greenberg, Roger  
CC TITLE OF INVENTION: Mouse Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 101  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/042,460  
CC FILING DATE: 16-MAR-1998  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,549  
CC FILING DATE: 19-NOV-1997

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,584  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/979,742  
CC FILING DATE: 26-NOV-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-003110US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3496 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MODELOC TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: -  
CC LOCATION: 1..3496  
CC OTHER INFORMATION: /note= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT) CDNA clone"  
CC FEATURE:  
CC NAME/KEY: misc\_feature  
CC LOCATION: 10..3435  
CC OTHER INFORMATION: /note= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT) CDNA"  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 39..3404  
CC OTHER INFORMATION: /product= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT)"  
CC SEQUENCE 3496 BP; 762 A; 998 C; 928 G; 808 T; 0 OTHER.

Query Match 29.0%; Score 825; DB 23; Length 3496;  
Best Local Similarity 67.1%; Pred. No. 0.00e+00;  
Matches 1901; Conservative 0; Mismatches 838; Indels 93; Gaps 5;

Db 23 CCGCGCTTGAGCACATGACCGCGCTCTCGTTGCCCGCGGTGCGCTCTCTGCTGCG 82  
QY 51 CCGCGCACACCCCGCGATGCGCGCGTCCCGCGTCCCGAGCGCTCCCTGCTCGG 110  
Db 83 CAGCCGATACCGGAGGTGCGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 142  
QY 111 CAGCCACTACCGGAGGTGCTGCGCTGGCCACGTTGTCGCGGCGCTGGGCGCCAGGG 170  
Db 143 CAGCGGCTTTGCAACCCCGGAGACCCGGAAGATCTACCGCACATTTGGTTGCCCAATGCT 202  
QY 171 CTGCGGCTGTGTCAGCGCGGAGACCCGCGGCTTCGCGCGCTGTGGCCAGTGCCT 230  
Db 203 AGTGTGATGACCTGGGGCTCAGAGCTCCACCTGCGGACCTTTCTTCACCAAGGTGC 262  
QY 231 GGTGTGCTGCTGGGACGACGCGCGCCCGCGCGCTCCCTTCCGCCAGGTGC 290  
Db 263 ATCCCTGAAGAGCTGGTGGCCAGGTTGTCAGAGACTCTGCGAGCGCAACGAGAGAA 322  
QY 291 CTGCTGAAGAGCTGGTGGCCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGGCGGAAGA 350  
Db 323 CGTGTGCTTTTGGCTTTGAGCTGCTTAACGAGCGCACGAGCGCGCTTCCCATGGCCTT 382  
QY 351 CGTGTGCTTTCGCTTCGCGCTGCTGGAGCGGCGCGCGCGGCGCGCGGCGCTT 410  
Db 383 CACTAGTAGCGTGGCTAGCTACTTGTCCCAACACACTGTTATTGAGACCTGCGTGCAGTGG 442  
QY 411 CACCACGACGCTGGCGAGCTACCTGCCCAACACACGCTGACCGACGCACTGCGGGGAGCGG 470  
Db 443 TGCATGGATGCTACTGTTGAGCCGAGTGGGCGAGCACTGCTGTCTACCTGCTGGGACA 502  
QY 471 GCGGTGGGGGCTGCTGTGCGCGCGCGCTGGCGAGCACTGCTGTTCACCTGCTGGCAGC 530





QY 2655 CTGCGTTTGGTGATTTCTTCTTGGTGACACCTCACCTCACCCACGCGAAGAACCTT 2714  
Db 2666 COTAGACACCTGGTCCATGCGGTTCCCTGAGATGGGTGCATGATAAAGTTCAGAGAC 2725  
QY 2715 CCTCAGGACCTGGTCCGAGGTGTCCTCTGAGTATGGCTGCGTGGTGAACCTTGGCGAAGAC 2774  
Db 2726 AGTGGTGAACCTCCCTGAGGACCTGTGTACCTGGTGGTGCAGCTCCATACCACTGCC 2785  
QY 2775 AGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGCAGGCTTTTGTTCAGATGCC 2834  
Db 2786 TGCTCAGTCGCT 2797  
QY 2835 GGCCACGCGCT 2846

## RESULT 14

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.  
AC xxxxxx  
DT  
DE Sequence 36, Application US/09026981  
CC Sequence 36, Application US/09026981  
CC GENERAL INFORMATION:  
CC APPLICANT: Counters, Christopher M.  
CC APPLICANT: Meyerson, Matthew  
CC APPLICANT: Weinberg, Robert A.  
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
CC STREET: Two Militia Drive  
CC CITY: Lexington  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02173  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/026,981  
CC FILING DATE: 20-FEB-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/064,322  
CC FILING DATE: 30-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/055,762  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/054,549  
CC FILING DATE: 01-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/047,151  
CC FILING DATE: 20-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/038,750  
CC FILING DATE: 20-FEB-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Granahan, Patricia  
CC REGISTRATION NUMBER: 32,227  
CC REFERENCE/DOCKET INFORMATION: WHI97-11p4AM  
CC TELEPHONE: 781-861-6240  
CC TELEFAX: 781-861-9540  
CC INFORMATION FOR SEQ ID NO: 36:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3346 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 28.7%; Score 816; DB 23; Length 3346;  
Best Local Similarity 84.9%; Pred. No. 0.00e+00;  
Matches 1027; Conservative 0; Mismatches 0; Indels 182; Gaps 1;  
Db 971 GGGTTGGCTGTTCGGGCCGACAGACACCTCTGCGTGAGAGATCTCTGCCAAGTTCC 1030  
QY 1640 GGGTTGGCTGTTCGGGCCGACAGACACCTCTGCGTGAGAGATCTCTGCCAAGTTCC 1699  
Db 1031 TGCACCTGGCTGATGAGTGTGTACGTCGAGCTGCTCAGGTCTTCTTTTATGTACGG 1090  
QY 1700 TGCACCTGGCTGATGAGTGTGTACGTCGAGCTGCTCAGGTCTTCTTTTATGTACGG 1759  
Db 1091 AGACCACGTTTCAAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGC 1150  
QY 1760 AGACCACGTTTCAAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGC 1819  
Db 1151 AAAGCATTGGAAATCAGACAGCAGCTTGAAGGGGTGACGTCGCGGAGCTGCGGAAGCAG 1210  
QY 1820 AAAGCATTGGAAATCAGACAGCAGCTTGAAGGGGTGACGTCGCGGAGCTGCGGAAGCAG 1879  
Db 1211 AGGTACGACGATCGGGAAGCCAGGCCCTCTGCTGACGTCAGAGCTCCGCTTCATCC 1270  
QY 1880 AGGTACGACGATCGGGAAGCCAGGCCCTCTGCTGACGTCAGAGCTCCGCTTCATCC 1939  
Db 1271 CCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGGACTAGCTGCTGGAGCCAGAACGT 1330  
QY 1940 CCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGGACTAGCTGCTGGAGCCAGAACGT 1999  
Db 1331 TCCGAGAGAAAAGAGGCCGCGGCTCTCAGCTCGAGGGTGAAGCAGCTGTTACGCGTGC 1390  
QY 2000 TCCGAGAGAAAAGAGGCCGCGGCTCTCAGCTCGAGGGTGAAGCAGCTGTTACGCGTGC 2059  
Db 1391 TCAACTACGAGCGGGCGCGGCCCTCTCTGGGCGCTCTGTGCTGGGCTTGGAGC 1450  
QY 2060 TCAACTACGAGCGGGCGCGGCCCTCTCTGGGCGCTCTGTGCTGGGCTTGGAGC 2119  
Db 1451 ATATCCAGAGGCTGGCGACCTTCTGCTGCTGCTGGGCGCCAGAACCCGCGCCTG 1510  
QY 2120 ATATCCAGAGGCTGGCGACCTTCTGCTGCTGCTGGGCGCCAGAACCCGCGCCTG 2179  
Db 1511 AGCTGACTTGTCAAGGTGATGTACGGGCGGTAGACACCATCCCGAGGACAGGC 1570  
QY 2180 AGCTGACTTGTCAAGGTGATGTACGGGCGGTAGACACCATCCCGAGGACAGGC 2239  
Db 1571 TCAGGAGGTGATCGCCAGCATCATCAAAACCCAGAACACACTGCTGCTGGCTCGGTATG 1630  
QY 2240 TCAGGAGGTGATCGCCAGCATCATCAAAACCCAGAACACACTGCTGCTGGCTCGGTATG 2299  
Db 1631 CCGTGGTCCAGAAAGCCCGCCATGGGACGTCGCGAAGGCTTCAAGAGCCACGTC---- 1686  
QY 2300 CCGTGGTCCAGAAAGCCCGCCATGGGACGTCGCGAAGGCTTCAAGAGCCACGTCCTA 2359  
Db 1686 ----- 1686  
QY 2360 CTTGACAGACCTTCAGCGCTACATGCGACAGTTCTGCTGGCTACCTGCAGGAGACAGCC 2419  
Db 1686 ----- 1686  
QY 2420 CGCTGAGGATGCCGCTCTCATCGAGCAGAGCTCTCTCCTGATGAGGCCAGCAGTGCC 2479  
Db 1687 -----CT 1688  
QY 2480 TCTTCGAGCTCTTCTACGCTTCTATGTGCCACACGCGCTGCGCATCAGGGCAAGTCT 2539  
Db 1689 AGTCCAGTCCAGGGGATCCCGAGGCTCTCATCTCTCCAGGCTGCTCTGACGCTGT 1748  
QY 2540 AGTCCAGTCCAGGGGATCCCGAGGCTCTCATCTCTCCAGGCTGCTCTGACGCTGT 2599  
Db 1749 GCTACGGGACATGGAGAACAAAGCTGTTTGGGGATTTGGCGGACGCGTGTCTCTGC 1808  
QY 2600 GCTACGGGACATGGAGAACAAAGCTGTTTGGGGATTTGGCGGACGCGTGTCTCTGC 2659  
Db 1809 GTTTGGTGATGATTTCTTGTGTTGACACCTCACCTCACCCACGCGAAGAACCTTCTCTCA 1868

|||||  
QY 2660 GTTGTGGATGATTTCTTTGTTGGTGACACCTCACCTCACCGGAAACCTTCTCA 2719  
Db 1869 GGACCTGTGTCGAGGTGCTCCTGAGTATGGCTGGTGAACCTTGCAGAACAGAGTGG 1928  
QY 2720 GGACCTGTGTCGAGGTGCTCCTGAGTATGGCTGGTGAACCTTGCAGAACAGAGTGG 2779  
Db 1929 TGAACCTTCCCTGTAGAACAGAGGCGCTGGTGGGACGCGCTTTTGTACAGATGCGCGGCC 1988  
QY 2780 TGAACCTTCCCTGTAGAACAGAGGCGCTGGTGGGACGCGCTTTTGTACAGATGCGCGGCC 2839  
Db 1989 ACGGCCTAT 1997  
QY 2840 ACGGCCTAT 2848  
|||||

RESULT 15  
ID US-08-912-951-3 STANDARD; DNA; UNC; 2176 BP.  
AC xxxxx  
DT  
Sequence 3, Application US/08912951.  
CC Sequence 3, Application US/08912951.  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITLE OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2176 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cdna  
CC FEATURE:  
CC NAME/KEY: -  
CC LOCATION: 1..2176  
CC OTHER INFORMATION: /note= "clone 712562"  
SQ SEQUENCE 2176 BP; 432 A; 678 C; 642 G; 422 T; 2 OTHER.  
  
Query Match 26.9%; Score 767; DB 22; Length 2176;  
Best Local Similarity 84.3%; Pred. No. 0.00e+00;  
Matches 978; Conservative 0; Mismatches 0; Indels 182; Gaps 1;  
  
Db 1 GCCCAAGTTCCTGCACCTGGCTGATGAGTGTGTACGTGCTCGAGCTGCTCAGGTCTTTCTT 60  
QY 1689 GCCCAAGTTCCTGCACCTGGCTGATGAGTGTGTACGTGCTCGAGCTGCTCAGGTCTTTCTT 1748  
Db 61 TTATGTACAGGAGACCCAGCTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTG 120  
QY 1749 TTATGTACAGGAGACCCAGCTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTG 1808  
Db 121 GAGCAAGTTGCAAGACATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCCGGAGCT 180  
QY 1809 GAGCAAGTTGCAAGACATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCCGGAGCT 1868  
Db 181 GTCGGAAGCAGAGGTACGGCAGCATCGGGAAGCCAGGCCCTCTGCTGAGCTCCAGACT 240  
QY 1869 GTCGGAAGCAGAGGTACGGCAGCATCGGGAAGCCAGGCCCTCTGCTGAGCTCCAGACT 1928  
Db 241 CGCTTCATCCCAAGCCTGACGGGCTCGGCCCGCATGTGTAACATGAGTACGTGCTGGG 300  
QY 1929 CGCTTCATCCCAAGCCTGACGGGCTCGGCCCGCATGTGTAACATGAGTACGTGCTGGG 1988  
Db 301 AGCCAGAACGTTCCGCAGAGAAAGAGGGCCGAGCGCTCTCACCTCGAGGGTGAAGGCACT 360  
QY 1989 AGCCAGAACGTTCCGCAGAGAAAGAGGGCCGAGCGCTCTCACCTCGAGGGTGAAGGCACT 2048  
Db 361 GTTACGCTGTCTCAACTACGAGCGGGCCGCGGCCCTCTCTGCGCGCTCTGTGCT 420  
QY 2049 GTTACGCTGTCTCAACTACGAGCGGGCCGCGGCCCTCTCTGCGCGCTCTGTGCT 2108  
Db 421 GGGCTTGACCATATCCACAGGGCTCGCGCACCTTCGTGCTGCGGTGCGGGGCCAGGA 480  
QY 2109 GGGCTTGACCATATCCACAGGGCTCGCGCACCTTCGTGCTGCGGTGCGGGGCCAGGA 2168  
Db 481 CCGCGCGCTGAGCTGTACTTTGTCAAGTGGATGTGACGGGCGCGGTACGACACCATCCC 540  
QY 2169 CCGCGCGCTGAGCTGTACTTTGTCAAGTGGATGTGACGGGCGCGGTACGACACCATCCC 2228  
Db 541 CCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAGTACTGCGT 600  
QY 2229 CCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAACAGTACTGCGT 2288  
Db 601 GCGTCGGTATGCCGTGGTCCAGAGAGGGCCGCCATGGGACGCTCCCGAAGGCGCTTCAAGAG 660  
QY 2289 GCGTCGGTATGCCGTGGTCCAGAGAGGGCCGCCATGGGACGCTCCCGAAGGCGCTTCAAGAG 2348  
Db 661 CCACGTC----- 667  
QY 2349 CCACGTCCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCA 2408  
Db 667 ----- 667  
QY 2409 GGAGACAGCGCGCTGAGGGATGCCGTGCTCATCGAGCAGAGCTCCTCCCTGAATGAGGC 2468  
Db 667 ----- 667



QY 2469 CAGCAGTGGCTCTTCACGCTTCCTACGCTTCATGTGCCACACGCGGTGCCATCAG 2528  
Db 668 -----CTACGTCAGTGGCCAGGGGATCCCGCAGGGTCCATCCTCTCCACGCTGCT 718  
QY 2529 GGGCAAGTCTCTACGTCCTCAGTGGCCAGGGGATCCCGCAGGGTCCATCCTCTCCACGCTGCT 2588  
Db 719 CTGCAGCCTGTGTACGCGGACATGGAGAACAAAGCTGTTGCGGGGATTCGGCGGGACGG 778  
QY 2589 CTGCAGCCTGTGTACGCGGACATGGAGAACAAAGCTGTTGCGGGGATTCGGCGGGACGG 2648  
Db 779 GCTGCTCTCGGTTTGGTGGATGATTTCTTGTGTCACACCTCACCTCACCCACGCGAA 838  
QY 2649 GCTGCTCTCGGTTTGGTGGATGATTTCTTGTGTCACACCTCACCTCACCCACGCGAA 2708  
Db 839 AACCTTCTCAGGACCTGTGTCAGGTGTCCTGTAGAGACGAGGCCCTGGGTGGCAGGCTTTTGTCA 898  
QY 2709 AACCTTCTCAGGACCTGTGTCAGGTGTCCTGTAGAGACGAGGCCCTGGGTGGCAGGCTTTTGTCA 2768  
Db 899 GAAGACAGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGCAGGCTTTTGTCA 958  
QY 2769 GAAGACAGTGGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGCAGGCTTTTGTCA 2828  
Db 959 GATCGCGGCCACGCGCTAT 978  
QY 2829 GATCGCGGCCACGCGCTAT 2848

Search completed: Fri Dec 25 17:47:26 1998  
Job time : 9031 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*

W P S R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 15:11:23 1998; MasPar time 245.67 Seconds  
Tabular output not generated. 540.766 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1-2848) from US08951733.seq (1 of 3)  
Perfect Score: 2848  
N.A. Sequence: 1 CACGCGTCCGGCGAGCGCTG.....GATCGCGGCCACGGCCTAT 2848  
Comp: GTGCGCAGGCCGCTCGGCAC.....CTAGCGCGGGTGGCGGATA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 9.326; Variance 5.654; scale 1.649

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	67	2.4	7218	1	US-08-232- Sequence 14, Applicati	2.08e-25
2	49	1.7	7218	1	US-08-232- Sequence 14, Applicati	1.91e-14
3	34	1.2	215	1	US-08-238- Sequence 5, Applicatio	6.10e-06
4	35	1.2	215	1	US-08-238- Sequence 5, Applicatio	1.76e-06
5	25	0.9	66	1	US-08-471- Sequence 144, Applicat	2.32e-01
6	25	0.9	68	1	US-07-977- Sequence 243, Applicat	2.32e-01
7	25	0.9	69	1	US-08-471- Sequence 142, Applicat	2.32e-01
8	26	0.9	74	2	PCT-US95-1 Sequence 94, Applicat	7.68e-02
9	25	0.9	74	2	PCT-US95-1 Sequence 100, Applicat	2.32e-01
10	26	0.9	74	2	PCT-US95-1 Sequence 100, Applicat	7.68e-02
11	26	0.9	74	2	PCT-US95-1 Sequence 99, Applicat	7.68e-02
12	25	0.9	75	2	PCT-US95-1 Sequence 99, Applicat	2.32e-01
13	25	0.9	75	2	PCT-US95-1 Sequence 99, Applicat	2.32e-01
14	26	0.9	81	2	PCT-US95-1 Sequence 92, Applicati	7.68e-02
15	25	0.9	81	2	PCT-US95-1 Sequence 98, Applicati	2.32e-01
16	26	0.9	81	2	PCT-US95-1 Sequence 98, Applicati	7.68e-02
17	26	0.9	81	2	PCT-US95-1 Sequence 92, Applicati	7.68e-02
18	26	0.9	82	2	PCT-US95-1 Sequence 97, Applicati	7.68e-02
19	26	0.9	82	2	PCT-US95-1 Sequence 97, Applicati	7.68e-02

20	25	0.9	242	1	US-08-273- Sequence 1, Applicatio	2.32e-01
c	21	0.9	1004	2	PCT-US95-0 Sequence 7, Applicatio	2.32e-01
c	22	0.9	1386	2	PCT-US95-0 Sequence 2, Applicatio	2.32e-01
c	23	0.9	1611	2	PCT-US93-0 Sequence 3, Applicatio	2.32e-01
c	24	0.8	59	2	PCT-US95-1 Sequence 95, Applicati	1.97e+00
c	25	0.8	65	1	US-08-471- Sequence 145, Applicat	6.84e-01
c	26	0.8	66	2	PCT-US95-1 Sequence 93, Applicati	6.84e-01
c	27	0.8	66	1	US-08-471- Sequence 144, Applicat	1.97e+00
c	28	0.8	68	1	US-08-471- Sequence 143, Applicat	6.84e-01
c	29	0.8	69	1	US-08-471- Sequence 142, Applicat	1.97e+00
c	30	0.8	84	2	PCT-US95-1 Sequence 25, Applicati	1.97e+00
c	31	0.8	84	2	PCT-US94-0 Sequence 25, Applicati	1.97e+00
c	32	0.8	84	1	US-08-133- Sequence 120, Applicat	1.97e+00
c	33	0.8	84	1	US-08-300- Sequence 25, Applicati	1.97e+00
c	34	0.8	225	1	US-08-299- Sequence 18, Applicati	1.97e+00
c	35	0.8	242	1	US-08-273- Sequence 1, Applicatio	6.84e-01
c	36	0.8	1404	3	5480796-8 Patent No. 5480796	1.97e+00
c	37	0.8	1404	3	5171840-8 Patent No. 5171840	1.97e+00
c	38	0.8	1640	1	US-07-807- Sequence 11, Applicati	1.97e+00
c	39	0.8	1640	1	US-08-299- Sequence 11, Applicati	1.97e+00
c	40	0.8	2061	3	5171840-1 Patent No. 5171840	1.97e+00
c	41	0.8	2061	3	5480796-1 Patent No. 5480796	1.97e+00
c	42	0.8	2253	3	5457037-2 Patent No. 5457037	1.97e+00
c	43	0.8	2454	1	US-07-872- Sequence 1, Applicatio	6.84e-01
c	44	0.8	2625	3	5457037-4 Patent No. 5457037	1.97e+00
c	45	0.8	3336	3	5457037-1 Patent No. 5457037	1.97e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx

Sequence 14, Application US/08232463  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA

ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:



CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/238,163  
CC FILING DATE: 03-MAY-1994  
CC CLASSIFICATION: 800  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Bastian, Kevin L.  
CC REGISTRATION NUMBER: 34,774  
CC REFERENCE/DOCKET NUMBER: 2307E-540  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 543-9600  
CC TELEFAX: (415) 543-5043  
CC INFORMATION FOR SEQ ID NO: 5:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 215 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: misc\_feature  
CC LOCATION: 1..215  
CC OTHER INFORMATION: /standard\_name= "Deduced amino acid  
CC OTHER INFORMATION: sequence of PGIP from bean."  
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.2%; Score 34; DB 1; Length 215;  
Best Local Similarity 12.9%; Pred. No. 6.10e-06;  
Matches 17; Conservative 56; Mismatches 59; Indels 0; Gaps 0;

Db 83 HTNVSGADSKTVTDYNSAGTSSNGGTGDNRSAGDSYSSKTAAMTSRRTGKTANNAYD 142  
QY 130 CTGCGCTGGCCAGCTTCGTCGGCGCTGGGCCCCAGGCTGGCGTGGTGCACGCG 189  
Db 143 SRNMGDASVSGDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGTGTSNVNNGGNGKRDVS 202  
QY 190 GGGGACCCGGGGCTTTCCGGCGCTGTGTGCGCCAGTGCCTGGTGGTGGCTGGGAC 249  
Db 203 SYANKCCGSSC 214  
QY 250 GCACGGCGGCC 261

RESULT 4  
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.  
AC xxxxxx  
DT Sequence 5, Application US/08238163  
DE Sequence 5, Application US/08238163  
CC Patent No. 556830  
CC GENERAL INFORMATION:  
CC APPLICANT: BENNETT, Alan  
CC APPLICANT: LABAVITCH, John M.  
CC APPLICANT: POWELL, Ann  
CC APPLICANT: STOFZ, Henrik  
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL  
CC TITLE OF INVENTION: POLYLACTONASES AND THEIR USE TO CONTROL FUNGAL DISEASE  
CC NUMBER OF SEQUENCES: 24  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESS: Townsend and Townsend Kourie and Crew  
CC STREET: Stewart Street Tower, One Market Plaza  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: US  
CC ZIP: 94105-1493  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 03-MAY-1994  
CC APPLICATION NUMBER: US/08/238,163  
CC CLASSIFICATION: 800

CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Bastian, Kevin L.  
CC REGISTRATION NUMBER: 34,774  
CC REFERENCE/DOCKET NUMBER: 2307E-540  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 543-9600  
CC TELEFAX: (415) 543-5043  
CC INFORMATION FOR SEQ ID NO: 5:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 215 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: misc\_feature  
CC LOCATION: 1..215  
CC OTHER INFORMATION: /standard\_name= "Deduced amino acid  
CC OTHER INFORMATION: sequence of PGIP from bean."  
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.2%; Score 35; DB 1; Length 215;  
Best Local Similarity 13.2%; Pred. No. 1.76e-06;  
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRFASNDKAKKDGNTTSWTTDCNRTWGVCDTFTYRVNDSGHNKYSAN 65  
Cp 1615 AGTCCCGACGCTCATCTCCAGCTCAGCTCCTGCAGCAGAGAGCTTGGCATGCTTCCCA 1556  
Db 66 YNTGGNNVGAATHTYHTNVSGADSKTVTDYNSAGTSSNGGTGDNRSAGDSYSSK 125  
Cp 1555 GGGAGATGAATCTTCTGGTGTCTTCTGAGGAGGCGGCTTCTGTGCTGGAGCCCA 1496  
Db 126 AMTSRRTGTANNADVSRNMGDASVSGDRNKKHAKNSADGKVGSKNNGDRNNRYGTGT 185  
Cp 1495 GGCTGGGGCACCAGCGCGGCGAGCGAGCGCCGACAGCCGCTACCTGCGAGGGGC 1436  
Db 186 KSNVSNNGGNGKRDVS 202  
Cp 1435 TGCTGTGCTGGCGGAGC 1419

RESULT 5  
ID US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.  
AC xxxxxx  
DT Sequence 144, Application US/08471052A  
DE Sequence 144, Application US/08471052A  
CC Patent No. 5625033  
CC GENERAL INFORMATION:  
CC APPLICANT: Kay, B. K.  
CC APPLICANT: Fowlkes, D. M.  
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
CC NUMBER OF SEQUENCES: 166  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036-2711  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/471,052A  
CC FILING DATE: 06-JUNE-1995  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872

[illegible]

```
Sequence 100, Application PC/TUS9511934  
Sequence 100, Application PC/TUS9511934
```

CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries

CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas

CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM.

CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/11934  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18 872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC







```

CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.

Query Match      0.9%; Score 25; DB 2; Length 81;
Best Local Similarity 7.7%; Pred. No. 2.32e-01;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db 13 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 72
Qy 945 CCACGCCACCCATCGGTGGCGCCGACGACGCGGGCCCCCATCCACATCGGGCC 1004
      : : : : : : : : : : : : : : : : : : : : : :
Db 73 ACCAC 77
      |||||
Qy 1005 ACCAC 1009

```

Search completed: Fri Dec 25 15:16:36 1998  
 Job time : 313 secs.

\*\*\*\*\*  
M P S R E L H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 09:43:38 1998; MagPar time 5827.06 Seconds  
Tabular output not generated.  
726.145 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1-1920) from US08951733.seq (1 of 2)  
Perfect Score: 1920  
N.A. Sequence: 1 CAGCGCGTCCGGCAGGCGTG.....CCAGGCGCGCGCTGCTGACG 1920  
Comp: GTCCGCGAGCGCGCTGCGGAC.....GGTCCGGCGGCGGACGACTGC

Scoring table:  
Gap open 30; Gap extend 1  
TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_lo 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_vl

Database: genbank107  
15:gb\_ba1 16:gb\_ba2 17:gb\_htg 18:gb\_in 19:gb\_om 20:gb\_ov  
21:gb\_pat 22:gb\_ph 23:gb\_pl1 24:gb\_pl2 25:gb\_pr1  
26:gb\_pr2 27:gb\_pr3 28:gb\_ro 29:gb\_st 30:gb\_sts 31:gb\_sy  
32:gb\_un 33:gb\_vl

Statistics: Mean 11.310; Variance 6.127; scale 1.846

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	1909	99.4	4015 27 AF015950 Homo sapiens telomeras 0.00e+00
2	1896	98.8	8960 31 AF043739 Synthetic construct hu 0.00e+00
3	1895	98.7	4027 28 AF018167 Homo sapiens telomeras 0.00e+00
4	453	23.6	3426 27 AF051911 Mus musculus telomeras 0.00e+00
5	451	23.5	3369 28 AF073311 Mus musculus telomeras 0.00e+00
6	110	5.7	201 28 AF029235 Mus musculus telomeras 3.62e-55
7	67	3.5	7218 21 I66494 Sequence 14 from paten 6.05e-25
8	44	2.3	7218 21 I66494 Sequence 14 from paten 4.31e-10
9	37	1.9	74371 26 AC005369 Homo sapiens chromosom 5.98e-06
10	34	1.8	215 21 I28278 Sequence 5 from patent 2.93e-04
11	35	1.8	215 21 I28278 Sequence 5 from patent 8.13e-05
12	34	1.8	74371 26 AC005369 Homo sapiens chromosom 2.93e-04

13	30	1.6	201 21	A10158	S.griseus gene for pre	4.16e-02
14	30	1.6	201 21	A10161	Synthetic DNA for prep	4.16e-02
c 15	30	1.6	201 21	A10162	Synthetic DNA for prep	4.16e-02
c 16	30	1.6	201 21	A10159	S.griseus gene for pre	4.16e-02
17	30	1.6	201 21	E04076	gDNA encoding envelope	4.16e-02
18	31	1.6	216021 26	HUAC004787	Homo sapiens Chromosom	1.24e-02
19	28	1.5	1663 19	MYU92534	Mustela vison microsai	4.40e-01
20	29	1.5	6011 33	HS4RSIR22T	Epstein-Barr virus (wi	1.37e-01
c 21	28	1.5	31812 24	SPUNK4	S.pombe chromosome I c	4.40e-01
c 22	28	1.5	39073 23	SPAC2E11	S.pombe chromosome I c	4.40e-01
c 23	29	1.5	76094 25	HS179D3A	Human DNA sequence fro	1.37e-01
c 24	27	1.4	565 21	E04076	gDNA encoding envelope	1.38e+00
c 25	27	1.4	904 33	HS4BHLF	Epstein-Barr virus Bam	1.38e+00
26	26	1.4	1341 20	CHKDILRB	Gallus domesticus DI-1	4.23e+00
27	26	1.4	2926 15	SMAPROBA	S.marcescens proBA ope	4.23e+00
28	26	1.4	2926 15	SMPROBAG	S.marcescens proBA gen	4.23e+00
29	26	1.4	4767 28	AF034863	Rattus norvegicus syna	4.23e+00
30	27	1.4	5050 33	HS4U2IR2A	Epstein-Barr virus (B9	1.38e+00
31	27	1.4	5400 15	SCACTVA	S.coelicolor 6 activa r	1.38e+00
32	26	1.4	12949 16	AF011922	Pseudomonas aeruginosa	4.23e+00
c 33	27	1.4	35980 16	MTY20H10	Mycobacterium tubercul	1.38e+00
c 34	26	1.4	43034 26	HSN5H6	Human DNA sequence fro	4.23e+00
35	27	1.4	64078 18	AC004365	Drosophila melanogaste	1.38e+00
c 36	26	1.4	85132 26	U82668	Homo sapiens shox gene	4.23e+00
37	26	1.4	123491 17	HS126A5	Human DNA sequence ***	4.23e+00
38	26	1.4	125815 17	HS436C18	Human DNA sequence ***	4.23e+00
39	26	1.4	138706 17	AC004946	*** SEQUENCING IN PROG	4.23e+00
40	26	1.4	147216 27	AF093117	Homo sapiens chromosom	4.23e+00
41	26	1.4	147542 17	HS754E20	Human DNA sequence ***	4.23e+00
42	27	1.4	172281 33	EBV	Epstein-Barr virus (EB	1.38e+00
43	27	1.4	184113 33	HS4B958RAJ	Epstein-Barr virus, ar	1.38e+00
44	25	1.3	69 21	I41362	Sequence 142 from pate	1.26e+01
45	25	1.3	201 21	A10159	S.griseus gene for pre	1.26e+01

ALIGNMENTS

RESULT	1	AF015950	4015 bp	mrna	PRI	16-AUG-1997
LOCUS						
DEFINITION						
ACCESSION						
NID						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
gene						
CDS						

/product="telomerase reverse transcriptase"

/product="telomerase reverse transcriptase"  
 /db\_xref="PID:g2330017"  
 /translation="MPAPRCRAVRSLLRSHYREVLPATFVRLGPQGRNLYVQRGDP  
 AAFRLVAQCLVCVMDRPPPPAPSRFRQVSLKELVRLVGLQCLCEGAKNVAJAFGA  
 LLDGARGGPEAFETTSVRSYLPNTVDLADSGKAWGLLLRRVDGDDLVLHLARLAFV  
 LVAPSCAYQVCGPLPYQLGAAATPPHAPSGRRLLCGRACRANHWSYREAGVFLGLPA  
 PGARRGGASRSGLSPYLRKPRGAAPPEPTPVQGGSWAHPRGTRSGDRGCVGVSPA  
 RPAAETSLBGLSPGTHSHPSYRGQHHAGPPTSRPDPWDTPCPVYAAETKHFLYS  
 SGDKOELRPSFLLSLSPSTGARRVLETFLGSRWMPMCTPRRLPQLPORYWQMRPL  
 FLELLGHAQCPGYVLLKTHCPRLAAVTPAAGVCAREKPGGSVAAREEDTDPRLVQ  
 LKQHSYPMQVYGVRAKLRRLYPGLGWSRHNRRFLRNKKFISLGKAKLSLQL  
 TLRQSVDRCAWLRSRPGVCGPAAEHLRREELIAKLHLWLSVYVVELLSYFTVET  
 TPKQNLRLFYVNSYKWSKLGQIGIRHLKRLQVRLSEAAVYRQREARPLLTLRLRLRI  
 PKPDGRLFYNNMDYVVGARTFREKRAERLTSRVKALFVLYNREARRPGLLGASVLG  
 LDIITHAARVFLVRRAQDPPEPLYFVKVDVTAIDTIPQDRLETVASIIKQNTYCL  
 VRYAVQKAAGHGVRAQPSHVSITLDLOPYMKRQVFAHLQETSPLDVAVIQBSNL  
 NEASSGLFDVFLREMHCHAVRIRGKSVQCQGIPOGSLSTLLCSLCYGMENKLPAG  
 IRDGLLRLVDLFLVYPLHTHAKTFLRLVLRGVPYEGCVNLRKTVNPPYDEDAL  
 GGTAVOMPAHGLFPMCLGDLTLEVDQSDYSYATSRSLTSRATFNFRGKAGNRMR  
 KLGRVLRKSLGFLDLQVNSLQVCTNIYKILLQVLRPHACVGLQFPFHQKQVNP  
 FFLRVSTADSLCSYKILLKNAQMSLGAKGAAGLPSEAVQWLCHQAKFLKLTSHRVT  
 YVPLLGSLRQAQTSKRLPGTITLTALAAANPALPSDEKTIID"  
 1363 a 1275 a 714 t  
 663 a

[illegible]

Query Match 99.4%; Score 1909; DB 27; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1909; Conservative 0; Mismatches 0; Indels 0;

Db	1	GCAGCGTGTGGTCTGTCTGCTGCACGTGGGAAGCCCTGGCGCCCGCCGACACCCCGCGATGCC	60
Qy	12	GCAGCGTGTGGTCTGTCTGCTGCACGTGGGAAGCCCTGGCGCCCGCCGACACCCCGCGATGCC	71
Db	61	GCGGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGCGCAGCACACTACGGGAGGTGCT	120
Qy	72	CGGGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGCGCAGCCACTACGGAGGTGCT	131
Db	121	GC CGCTGGCCACGTTCTGTGCGGGCCCTGGGGCCCCAGGGCTGCGGCTGGTGACGCGGG	180
Qy	132	GC CGCTGGCCACGTTCTGTGCGGGCCCTGGGGCCCCAGGGCTGCGGCTGGTGACGCGGG	191
Db	181	GGACCCGGCGGCTTTCGGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGCTGGGACGC	240
Qy	192	GGACCCGGCGGCTTTCGGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGCTGGGACGC	251
Db	241	ACGGCCGCCCCCGCGCCGCCCTCCTTCGCCCAGGTGTCCTGCGCTGAAGAGAGTGTGTGC	300
Qy	252	ACGGCGGCCCCCGCGCCGCCCTCCTTCGCCCAGGTGTCCTGCGCTGAAGAGAGTGTGTGC	311
Db	301	CCGAGTCTCGAGAGGCTGTGCGAGCGCGCGCGGAGAACGTGCTGGCTTGGGCTTTCGC	360
Qy	312	CCGAGTCTCGAGAGGCTGTGCGAGCGCGCGCGGAGAACGTGCTGGCTTGGGCTTTCGC	371
Db	361	GCTGCTGAGCGGGGCCCGCGGGGCCCCCCCGAGGGCTTCACCACACAGCGTGGCGACGTA	420
Qy	372	GCTGCTGAGCGGGGCCCGCGGGGCCCCCCCGAGGGCTTCACCACACAGCGTGGCGACGTA	431
Db	421	CCTGCCAACACAGTGTACCGACGACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGTGCG	480
Qy	432	CCTGCCAACACAGTGTACCGACGACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGTGCG	491
Db	481	CCGGTGGGGACGACGTGCTTCACTGCTTGGCAGCGTGGCGCTCTTTGTGCTGCT	540
Qy	492	CCGGTGGGGACGACGTGCTTCACTGCTTGGCAGCGTGGCGCTCTTTGTGCTGCT	551
Db	541	GGCTTCCAGCTTGGGCTTACAGAGTGTGCGGGCCCGCGCTGTACCGCTTGGCGCTGCCAC	600
Qy	552	GGCTTCCAGCTTGGGCTTACAGAGTGTGCGGGCCCGCGCTGTACCGCTTGGCGCTGCCAC	611
Db	601	TCAGGCCGGGCCCGCCACACGCTAGTGGACCCGAGCGGCTCTGGGATGCGAAGCGGC	660
Qy	612	TCAGGCCGGGCCCGCCACACGCTAGTGGACCCGAGCGGCTCTGGGATGCGAAGCGGC	671

|||||  
QY 1752 TGTCCAGGACACGCTTTCAAAGAACAGCTCTTTTCTACCGAGAGTGCTCGAG 1811  
|||||  
Db 1801 CAAGTTCAAAGCATTTGAATCAGACGACTTGAAGAGGGTGCAGCTGCGGGAGCTGTC 1860  
|||||  
QY 1812 CAAGTTCAAAGCATTTGAATCAGACGACTTGAAGAGGGTGCAGCTGCGGGAGCTGTC 1871  
|||||  
Db 1861 GGAAGCAGAGTTCAGGAGCATTCGGGAAGCAGCCCGCCCTGCTGAGC 1909  
|||||  
QY 1872 GGAAGCAGAGTTCAGGAGCATTCGGGAAGCAGCCCGCCCTGCTGAGC 1920  
|||||

RESULT 2 AF043739 8960 bp DNA SYN 05-APR-1998  
LOCUS  
DEFINITION Synthetic construct human telomerase catalytic subunit  
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete  
sequence.  
ACCESSION AF043739  
NID g3023054  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddie,S.D.,  
Haber,D.A. and Weinberg,R.A.  
TITLE Telomerase activity is restored in human cells by ectopic  
expression of hTERT (hEST2), the catalytic subunit of telomerase  
Oncogene 16 (9), 1217-1222 (1998)  
98187713  
2 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddie,S.D.,  
Haber,D. and Weinberg,R.A.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,  
Cambridge, MA 02142, USA

FEATURES  
source  
1. .8960  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
order(1. 1101,4606. 8960)  
/note="mammalian expression vector; see GenBank Accession  
Number U47120"  
/organism="Cloning vector pCI-neo"  
/db\_xref="taxon:45772"  
1102. .4542  
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.33"  
1102. .4542  
/gene="hTERT/hEST2"  
1147. .4605  
/note="telomerase/reverse transcriptase"  
/codon\_start=1  
/transl\_table=11  
/product="telomerase catalytic subunit"  
/db\_xref="PID:g3023055"  
/translation="MPRAPKRAVRSLRSHYREVLPLATFVRKLPQGNLVQRGDP  
AAFLVAQCLVCPWDPAPPPAPSPFQVSLKELVARVLQRLCERGAKNVLAFGFA  
LLDGARGPFAEFTTSVSYLPTNTVDALRGSGAWGLLRVGGDDVLVHLARCALEV  
LVAPSCAYQVCPFLYOLGAATQAPPPHAGSPRRRLGRCERAWNHSVREAGVPLGLPA  
PGARRGGSASRLPLPKPRRGAAPERTPVQGGSWAHPGRTGSPDRGFCVSPFA  
RPAEATSLSGLSTRSHSPVGRQHHAGPPSTSRPPWDTPCPVYAEATKHEFLS  
SGDREQLRPSFLSLSPSTGARRLVETIFLGRSPWMPGTPRRLPRLPQRYWQMRPL  
FLELHGHAQCPYGVLLKTHCPPLRAAVTPAAGVCAREKPGQSVAAPEEDTPRLVQ  
LLRQSHSPWQYQVLAQCLRLVPPGLWGRHNERFLRNTKKEIFSLGKHLQSLQEL  
TKMSVRGCAWLRRSPGVCPAAEHLREELAKFLHLMVSVYVVELLSFFVFTET  
TFQKRLFFYKRSVMSLQSIGIFQHLKRVQLRELSAEVROHREAPALLTSPLRFI  
PKDGLRPIVNDIVVGARTREKRAERLTSRVKALFSLVNIYERARRPGLLGASVLG  
LDDTHRAWRTFVLVRADQPPPELYFVKVDVTGAYDTIPQDRLTEVLAETIKPONTYC  
VRRYAVVQKAHGHVHRKAFKSHVSTLTLDLPYMRQFVAHLQETSPLRDAVVIEQSSSL

NEASSCLFDVFLRFMCHAVRIRGKSYVOCQIPOGSIILSTLLCSLCYGDMEKLFAG  
IRRDGLLLVDDFLVTPHLTHAKTFLRLTVRGVPEYICGVNLRITVNFVPEDEAL  
GTGAFQMPAHGLFPWCGLLDTRTLEVDSDSYARTSIRASLTFRNFKAGNNRR  
KLFVLRKCHSLFLDQVNSLTQVNTYKILLQAYRFHACVLAQLPFOQVKNRPT  
FELRVISDTASLCSYSLKAKNAGMSLGAAGPLPSEAVQWLCHQAFLLKLTFRHRT  
VYPLGLSLRTAQTLRSKRLPGTTLTAAEAANPALPSDFKTLILDYVDVDPYAGYPYD  
VPDYAG"

source 4543. .4605  
/note="engineered based on the epitope for the  
anti-Influenza virus hemagglutinin monoclonal antibody  
12CA5; similar to part of the Influenza A hemagglutinin  
gene sequence in GenBank Accession Number M29257"  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others  
ORIGIN

Query Match 98.8%; Score 1896; DB 31; Length 8960;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1898; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1101 CGTCCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCCGCGCTCC 1160  
|||||  
QY 21 CGTCCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCGACCCCGCGATGCCGCGCTCC 80  
|||||  
Db 1161 CCGCTGCCGAGCGCTGCGCTCCCTGTGCGCAGCCACTACCGCAGGTGTGCGCTGGC 1220  
|||||  
QY 81 CCGCTGCCGAGCGCTGCGCTCCCTGTGCGCAGCCACTACCGCAGGTGTGCGCTGGC 140  
|||||  
Db 1221 CACGTTGCTGGCGCGCTGGGCGCCAGGGCTGGCTGGTGACGCGCGGGACCGGC 1280  
|||||  
QY 141 CACGTTGCTGGCGCGCTGGGCGCCAGGGCTGGCTGGTGACGCGCGGGACCGGC 200  
|||||  
Db 1281 GGCTTTCCGCGCGCTGGTGCGCCAGTGCCTGGTGGTGGCCCTGGGACGACGCGCGCC 1340  
|||||  
QY 201 GGCTTTCCGCGCGCTGGTGCGCCAGTGCCTGGTGGTGGCCCTGGGACGACGCGCGCC 260  
|||||  
Db 1341 CCCCGCGCGCGCTTCCTTCGCGCAGGTGTCTGCTGAAAGAGTGTGTGGCCCCGAGTGCT 1400  
|||||  
QY 261 CCCCGCGCGCGCTTCCTTCGCGCAGGTGTCTGCTGAAAGAGTGTGTGGCCCCGAGTGCT 320  
|||||  
Db 1401 GCAGAGCTGTGCAGCGCGCGCGGAAGAGCTGCTGGCCTTCGGCTTCGCGCTGCTGGA 1460  
|||||  
QY 321 GCAGAGCTGTGCAGCGCGCGCGGAAGAGCTGCTGGCCTTCGGCTTCGCGCTGCTGGA 380  
|||||  
Db 1461 CGGGCGCGCGCGCGCGCGCGCGAGGCTTCACCCACAGCGTGCAGCTACCTGCCCAA 1520  
|||||  
QY 381 CGGGCGCGCGCGCGCGCGCGCGAGGCTTCACCCACAGCGTGCAGCTACCTGCCCAA 440  
|||||  
Db 1521 CACGGTGACCGACGACTGCGGGGAGCGGGGCTGTGGGGCTGTGTGGCCCCGCTGGG 1580  
|||||  
QY 441 CACGGTGACCGACGACTGCGGGGAGCGGGGCTGTGGGGCTGTGTGGCCCCGCTGGG 500  
|||||  
Db 1581 CGACGAGTGTGTTTACCTGTGCGACGCTGGCGCGCTTTTGTGTGGTGGTCCAG 1640  
|||||  
QY 501 CGACGAGTGTGTTTACCTGTGCGACGCTGGCGCGCTTTTGTGTGGTGGTCCAG 560  
|||||  
Db 1641 CTGCGCTTACGAGTGTGCGGGCGCGCTGTACAGCTCGCGCGCTGCACCTCAGGCGCG 1700  
|||||  
QY 561 CTGCGCTTACGAGTGTGCGGGCGCGCTGTACAGCTCGCGCGCTGCCACTCAGGCGCG 620  
|||||  
Db 1701 GCCCGCGCCACACGCTAGTGGACCCCGAAGCGCTCTGGGATCGGAACGGGCTTGAACCA 1760  
|||||  
QY 621 GCCCGCGCCACACGCTAGTGGACCCCGAAGCGCTCTGGGATCGGAACGGGCTTGAACCA 680  
|||||  
Db 1761 TAGGCTCAGGAGCGCGGGTCCCTCGCGCGCTGCCAGCCCCGGGTGCGAGAGCGCGGG 1820  
|||||  
QY 681 TAGGCTCAGGAGCGCGGGTCCCTCGCGCGCTGCCAGCCCCGGGTGCGAGAGCGCGGG 740  
|||||  
Db 1821 GGGCAGTCCAGCGAGTCTGCCGCTTCCCAAGAGCGCGCGCTGCGCTGCCCTGA 1880  
|||||  
QY 741 GGGCAGTCCAGCGAGTCTGCCGCTTCCCAAGAGCGCGCGCTGCGCTGCCCTGA 800  
|||||



[illegible]

D	b	1092	CTACTCAGCTCTCTGAGGCCACAGCTGACTGCGCTCGAGAGCTCGTGAGACCATCTTT	1151
Q	y	1102	CTACTCAGCTCTCTGAGGCCACAGCTGACTGCGCTCGAGAGCTCGTGAGACCATCTTT	1161
D	b	1152	CTGGTTTCAGGCGCTTGATGTCAGAGGACTCCCGCAGGTTGCCCGCTCGCCACAGCG	1211
Q	y	1162	CTGGTTTCAGGCGCTTGATGTCAGAGGACTCCCGCAGGTTGCCCGCTCGCCACAGCG	1221
D	b	1212	TACTGGCAATGCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCAGTGCCCTAC	1271
Q	y	1222	TACTGGCAATGCGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCAGTGCCCTAC	1281
D	b	1272	GGGTGCTCTCAAGACGACTGCCCGCTGCGAGTGGCGTCAACCCAGCAGCGGTGTC	1331
Q	y	1282	GGGTGCTCTCAAGACGACTGCCCGCTGCGAGTGGCGTCAACCCAGCAGCGGTGTC	1341
D	b	1332	TGTGCCGGGAGAACCCACAGGCTCTGTGGCGGCCCCCGAGGAGGAGACACAGACCC	1391
Q	y	1342	TGTGCCGGGAGAACCCACAGGCTCTGTGGCGGCCCCCGAGGAGGAGACACAGACCC	1401
D	b	1392	CGTGCCTGTGTGAGCTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTG	1451
Q	y	1402	CGTGCCTGTGTGAGCTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCTTCGTG	1461
D	b	1452	CGGCGCTGCTCGCGCGGTGTGTGCCCCAGGCGCTCTGGGGTTCGAGCACAAAGACGC	1511
Q	y	1462	CGGCGCTGCTCGCGCGGTGTGTGCCCCAGGCGCTCTGGGGTTCGAGCACAAAGACGC	1521
D	b	1512	CGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTG	1571
Q	y	1522	CGCTTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTG	1581
D	b	1572	CAGGAGTGAAGTGAAGTGAAGTGTGCGGGCTGCGTGGCTGCGCAGGAGCCAGGG	1631
Q	y	1582	CAGGAGTGAAGTGAAGTGAAGTGTGCGGGCTGCGTGGCTGCGCAGGAGCCAGGG	1641
D	b	1632	GTGGCTGTCTCCGGCCGACAGCACCGTCTGCTGAGGAGATCTTGCCCAAGTTCTCTG	1691
Q	y	1642	GTGGCTGTCTCCGGCCGACAGCACCGTCTGCTGAGGAGATCTTGCCCAAGTTCTCTG	1701
D	b	1692	CAGTGGCTGATGATGTGTACGCTCGAGCTGCTCAGGTCTTTCTTTATGTCAAGG	1751
Q	y	1702	CAGTGGCTGATGATGTGTACGCTCGAGCTGCTCAGGTCTTTCTTTATGTCAAGG	1761
D	b	1752	ACCAGCTTCCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA	1811
Q	y	1762	ACCAGCTTCCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA	1821
D	b	1812	AGCATTTGAATCAGACGACTTGAAGAGGTGTCAGCTGCGGAGCTGTCGGAAGCAGAG	1871
Q	y	1822	AGCATTTGAATCAGACGACTTGAAGAGGTGTCAGCTGCGGAGCTGTCGGAAGCAGAG	1881
D	b	1872	GTACGAGCAGTCCGAAGCCAGGCCCGCCCTGCTGACG	1910
Q	y	1882	GTACGAGCAGTCCGAAGCCAGGCCCGCCCTGCTGACG	1920
RESULT		4		
LOCUS		AF051911	3426 bp	mRNA
DEFINITION		Mus musculus telomerase reverse transcriptase mRNA, complete cds		02-APR-1998
ACCESSION		AF051911		
NID		g3005591		
KEYWORDS		house mouse.		
SOURCE		Mus musculus		
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE		1 (bases 1 to 3426)		
AUTHORS		Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.		
TITLE		Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation		
JOURNAL		Oncogene (1998) in press		
REFERENCE		2 (bases 1 to 3426)		

RESULT	4				
LOCUS	AF051911	3426 bp	mrna	ROD	02-APR-1998
DEFINITION	Mus musculus telomerase reverse transcriptase mRNA, complete cds.				
ACCESSION	AF051911				
NID	g3005591				
KEYWORDS	house mouse.				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 3426)				
AUTHORS	Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.				
TITLE	Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation				
JOURNAL	Oncogene (1998) In press				
REFERENCE	2 (bases 1 to 3426)				









Db 949 TCACACCCCGCCAAATGCCCTTCAGCTCAGGCCATTATTTGACACAGACATTTCCCTT 1008  
QY 1000 CGCGCCACCACTCCCTGGGACACGCTTGTCCCGGTGTACGCGGACCAAGCACTTC 1059  
Db 1009 TACTCCAGGGAGATGGCCAAAGAGCTCTAAACCCCTATTCTCTACTCAGCAACCTCCAG 1068  
QY 1060 CTCTACTCTCAGGGACAAAGAGAGCTGCGGCCCTCTCTCTACTCAGCTCTCTGAGG 1119  
Db 1069 CTAATACTGACTGGGCCAGGAGACTGTTGGAGATCATCTTTCTGGGTCTCAAGGCTAGG 1128  
QY 1120 CCCAGCCTGACTGGCGCTCGGAGGCTCGTGGAGACATCTTTCTGGTTCAGAGCCCTGG 1179  
Db 1129 ACATCAGGACCACTCTGCAGGACACACCGTCTATCGCTGCATATGCGAGATGCGGCC 1188  
QY 1180 ATGCGAGGACTCCCGCAGGTGGCCCGCTGCCAGCGCTACTGCAATATGCGGCC 1239  
Db 1189 CHTTCCAAAGCTGCTGCTGAACCATGAGAGTGCATATGTCAGACTCTCTCAGGTCA 1248  
QY 1240 CHGTTCTGGAGCTGCTTGGGAACACCGCGAGTGCCTCAGGGGTGCTCTCAGAGG 1299  
Db 1249 CATTCAGGTTTTCGACAGCAACCAACAGGTGACAGATGCTT----- 1292  
QY 1300 CACTGCCCGCTCGAGCTCGGCTCACCCAGCAGCGGCTGTGTGCCCGGAGAGGCC 1359  
Db 1293 -----GACACAGCCCAACCGCACTCATGATTG 1323  
QY 1360 CAGGCTCTGTGGCGCCCGCCGAGGAGGAGACACAGACCCCGCTGCTGTCAGGTG 1419  
Db 1324 CTCGCGCTCACAGAGTCCCTGGGAGGTATATGTTTCTTCGGCCGCTCTCTGCAAG 1383  
QY 1420 CTCGCGCACACAGACGCCCTTGGCAGGTGTACGGCTTCGTCGGGCCCTGCTCGCCGG 1479  
Db 1384 GTGGTCTCTGCTAGTCTTGGGTACAGGCACATAGCGCGCTTCTTTAAGAACTTA 1443  
QY 1480 CTGTTGCCCGCCAGGCTCTGGGCTCCAGGCACACGACGCGCTTCTCAGGACACC 1539  
Db 1444 AAGAAGTTATCTGTTGGGAAATACGCCAAGCTATCACTCGAGAACTGATGTGAAG 1503  
QY 1540 AAGAAGTTATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTCGAGAGCTGACGTGGA 1599  
Db 1504 ATGAAGTAGAGATTGCCACTGGCTCCGACAGCGCGGGAAGACCGTCTCCCGCT 1563  
QY 1600 ATGAGCGTCCGGGACTGCGCTTGGCTGCGCAGGAGCCAGGGGTGGTGTCTCCGCC 1659  
Db 1564 GCAGACACCGCTCTGAGGAGAGGATCTCTGCTAGCTTCTCTTCTGCTGATGGACACA 1623  
QY 1660 GCAGACACCGCTCTGCTGAGGAGATCTTGCCCAAGTCTCTGCTGCTGATGATGTG 1719  
Db 1624 TAGTGGTACAGCTGTTAGTCTATCTTTTACATCAGAGAGGACATTCAGAGAAGAC 1683  
QY 1720 TAGCTGCTGAGCTGCTCAGTCTTTCTTTTATGTACCGGAGACACGCTTCAAAGAAC 1779  
Db 1684 AGGCTCTCTCTACCGTAAGAGTGTGTGAGCAGAGCTCAGAGCATTTGAGTCAAGGAA 1743  
QY 1780 AGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTCGAAGCATTTGAATCAGACAG 1839  
Db 1744 CACCTTGAGAGTGGCGCTACGGGAGCTGTCTACAGAGGAGGCTCAGGACATCAGCAGGAC 1803  
QY 1840 CACTTGAAGAGGTCAGCTCGGGAGCTGTCTGGAAGCAGAGGTTCAGGAGCATCGGNA 1899  
Db 1804 ACTGTGCTAGCATGC 1819  
QY 1900 GCCAGGCGCGCCCTGC 1915

RESULT 6  
LOCUS AF029235 201 bp mRNA ROD 12-NOV-1997  
DEFINITION Mus musculus telomerase catalytic subunit mRNA, partial cds.  
ACCESSION AF029235  
NID 92605902  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Mus musculus

Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 201)  
AUTHORS Drissi,R. and Cleveland,J.L.  
TITLE Partial sequence of Mus musculus telomerase catalytic subunit homolog  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 201)  
AUTHORS Drissi,R. and Cleveland,J.L.  
TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1997) Biochemistry, St Jude Children's Research Hospital, 322 North Lauderdale, Memphis, TN 38105, USA  
FEATURES  
source 1..201  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_line="32D.3"  
/codon\_start=1  
/product="telomerase catalytic subunit"  
/db\_xref="PID:g2605903"  
/translation="FFVYTESTFKNRLFFYRKSVMKLSIGVHQHLRVRLRELQSO  
EEVYRHODTWLAMPICLREIPK"  
BASE COUNT 51 a 54 c 55 g 41 t  
ORIGIN  
Query Match 5.7%; Score 110; DB 28; Length 201;  
Best Local Similarity 82.0%; Pred. No. 3,62e-55;  
Matches 141; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
Db 1 TTTCTTTATGTCAGGAGAGCACAATCCAGAAACACAGGCTTCTTCTACCGTAAGAGT 60  
QY 1744 TTTCTTTATGTCAGGAGAGCACCAGCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGT 1803  
Db 61 GTGTGGGACCAAGCTGCAGAGCATTGGAGTCAGGCAACACCTTGAGAGAGTCGGCTACGG 120  
QY 1804 GTCTGGAGCAAGTGTCAAAGCATTGGAATCAGACAGCACTTGAAGAGGTCAGCTCGG 1863  
Db 121 GAGCTGTCAAGAGGAGGTGAGCAGTCAGGACACCTGAGTACGCTGAGCATGC 172  
QY 1864 GAGCTGTGAGGAGGAGGTGAGCAGTCAGGAGCATCGGAGCCGCGCTGC 1915  
RESULT 7  
LOCUS I66494 7218 bp DNA PAT 23-DEC-1997  
DEFINITION Sequence 14 from patent US 5670367.  
ACCESSION I66494  
NID 92724471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 7218)  
AUTHORS Dornier,F., Scheifflinger,F. and Falkner,F.Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES  
source 1..7218  
/organism="unknown"  
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others  
ORIGIN  
Query Match 3.5%; Score 67; DB 21; Length 7218;  
Best Local Similarity 2.4%; Pred. No. 6,05e-25;  
Matches 9; Conservative 211; Mismatches 153; Indels 0; Gaps 0;  
Db 1055 GGAGCTTCGATATTT 1114  
QY 921 GGAGGTCGCTCTCTGCGAGCGCCACTCCACCCATCCGTCGCGCCGACCAAGCGC 980  
Db 1115 YY 1174  
QY 981 GGGCCCCCATCCACATCGCGGCCACCACTCCCTCGGACACAGCTTGTCCCGGTGTA 1040









[illegible]

Search completed: Thu Dec 24 13:04:28 1998  
Job time : 12050 secs.

**THIS PAGE BLANK (USPTO)**



\*\*\*\*\*

W P E R E L F (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 13:04:47 1998; MasPar time 5213.74 Seconds  
Tabular output not generated. 659.468 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1-1920) from US08951733.seq (1 of 2)  
Perfect Score: 1920  
N.A. Sequence: 1 CAGCGCTCCGGCAGCGCTG.....CCAGCGCCGCCCTGCTGACG 1920  
Comp: GTGCGCAGCGCCGTCGGAC.....GGTCCGGCGGCGGACGACTGC

Scoring table:  
Gap open 30; Gap extend 1  
TABLE default  
Gap open 30; Query 0  
Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs; 895388244 bases x.2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est55  
Database: genbank-est107  
5:gb\_est1 2:em\_gss1 3:em\_gss2 4:em\_gss3  
10:gb\_est11 11:gb\_est12 12:gb\_est13 13:gb\_est14  
14:gb\_est15 15:gb\_est16 16:gb\_est17 17:gb\_est18  
18:gb\_est19 19:gb\_est20 20:gb\_est21 21:gb\_est22 22:gb\_est23  
23:gb\_est24 24:gb\_est25 25:gb\_est26 26:gb\_est27 27:gb\_est28  
28:gb\_gss3 29:gb\_gss4

Statistics: Mean 11.357; Variance 2.736; scale 4.151

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	227	11.8	389	8	AA281296 zt08g02.r1 NCI_CGAP_GC	0.00e+00
2	58	3.0	252	12	AA754459 97SN1787 Rice Immature	1.27e-46
3	54	2.8	252	12	AA754459 97SN1787 Rice Immature	7.79e-41
4	49	2.6	247	12	AA754458 97SN1784 Rice Immature	9.33e-34
5	44	2.3	247	12	AA754458 97SN1784 Rice Immature	6.93e-27
6	31	1.6	660	11	AF034177 Homo sapiens ntcon6 co	2.10e-10
7	28	1.5	1287	12	AF038250 Homo sapiens clone ntc	5.22e-07
8	28	1.5	1287	12	AF038250 Homo sapiens clone ntc	5.22e-07
9	27	1.4	317	27	A0008007 CIT-HSP-2288C3.TF CIT-	6.33e-06
10	26	1.4	400	16	H44707 YP24c08.r1 Homo sapien	7.23e-05
11	27	1.4	660	11	AF034177 Homo sapiens ntcon6 co	6.33e-06
12	26	1.4	2275	11	AF034173 Homo sapiens ntcon2 co	7.23e-05

C	13	25	1.3	143	18	AI171394	EST217350 Normalized r	7.73e-04
	14	25	1.3	181	6	AA386387	VS81369 prostate glan.	7.73e-04
	15	25	1.3	187	13	AA855630	w68h11.r1 Stratagene	7.73e-04
	16	25	1.3	213	16	R72797	Y109c09.r1 Homo sapien	7.73e-04
	17	25	1.3	242	15	AI012107	EST206558 Normalized r	7.73e-04
	18	25	1.3	278	13	AA882818	TENS0393 T. cruzi epim	7.73e-04
	19	25	1.3	302	15	AI010844	EST205295 Normalized r	7.73e-04
	20	25	1.3	318	16	H50134	yo27a07.r1 Homo sapien	7.73e-04
	21	25	1.3	336	18	AI172396	EST218402 Normalized r	7.73e-04
	22	25	1.3	359	24	AA303595	EST16305 Aorta endothe	7.73e-04
	23	25	1.3	382	5	R54656	YJ74d04.r1 Homo sapien	7.73e-04
	24	25	1.3	420	17	AI071447	UI-R-C1-ku-c-05-0-UI.s	7.73e-04
	25	25	1.3	438	14	AA924206	UI-R-11-dx-a-11-0-UI.s	7.73e-04
	26	25	1.3	478	14	AA923372	UI-R-11-ee-d-07-0-UI.s	7.73e-04
	27	25	1.3	487	16	R86860	ym86a03.r1 Homo sapien	7.73e-04
	28	25	1.3	504	15	AI010120	EST204571 Normalized r	7.73e-04
	29	25	1.3	511	15	AA997735	UI-R-C2-hn-b-08-0-UI.s	7.73e-04
	30	25	1.3	511	17	AI072026	UI-R-C2-nd-e-12-0-UI.s	7.73e-04
	31	25	1.3	540	14	AA901048	UI-R-11-dv-e-09-0-UI.s	7.73e-04
	32	25	1.3	590	10	AA622227	no3c05.sl NCI_CGAP.Pr	7.73e-04
	33	25	1.3	677	18	AI105428	EST214717 Normalized r	7.73e-04
	34	25	1.3	705	14	AA941159	LD25143.5prime LD Dros	7.73e-04
	35	24	1.2	275	14	AA899633	UI-R-20-dg-f-09-0-UI.s	7.71e-03
	36	24	1.2	288	23	AA077661	7B35F02 Chromosome 7 F	7.71e-03
	37	24	1.2	394	16	R92606	yq07g02.sl Homo sapien	7.71e-03
	38	24	1.2	398	5	TA7934	ybi8d04.sl Homo sapien	7.71e-03
	39	24	1.2	426	9	AA591041	vl47b11.sl Stratagene	7.71e-03
	40	24	1.2	456	6	C19919	Rice cDNA, partial seq	7.71e-03
	41	24	1.2	496	23	AA243831	zr67b09.r1 Soares NHM	7.71e-03
	42	24	1.2	496	10	AA429394	zw32h12.sl Soares ovar	7.71e-03
	43	24	1.2	550	7	AA495273	fa02e07.r1 Zebrafish i	7.71e-03
	44	24	1.2	607	18	AI170655	EST216590 Normalized r	7.71e-03
	45	24	1.2	2275	11	AF034173	Homo sapiens ntcon2 co	7.71e-03

ALIGNMENTS

RESULT 1  
LOCUS AA281296 389 bp mRNA EST 14-AUG-1997  
DEFINITION zt08g02.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:712562 5', mRNA sequence.  
ACCESSION AA281296  
NID g1924194  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;  
Homo.  
REFERENCE 1 (bases 1 to 389)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 2187 Std Error: 0.00  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 385.  
Location/Qualifiers  
1..389  
/organism="Homo sapiens"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD+),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 2187 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 385.

Location/Qualifiers

1..389

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD+),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer

```

[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:712562"
/clone_lib="NCL-CPAP-GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
<1..>389
BASE COUNT      87 a 102 c 123 g  77 t
ORIGIN

Query Match      11.8%; Score 227; DB 8; Length 389;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 60
|||||
Qy 1690 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 1749
|||||

Db 61 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 120
|||||
Qy 1750 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 1809
|||||

Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 180
|||||
Qy 1810 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 1869
|||||

Db 181 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 231
|||||
Qy 1870 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 1920
|||||

RESULT 2
LOCUS      AA754459      252 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION  AA754459
NID        92801165
KEYWORDS   EST.
SOURCE     rice.
ORGANISM   Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
REFERENCE  1 (bases 1 to 252)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)
COMMENT   Contact: Eun M.Y.
          Department of Cytogenetics
          National Inst. of Agri. Sci. and Tech, RDA
          Suwon, Kyungido, Korea
          Tel: 82 331 290 0301
          Fax: 82 331 290 0307
          Email: myeun@sun20.asti.re.kr
          Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
          University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
          Seq primer: M13 Reverse Primer.
          Unpublished (1998)

FEATURES
source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"

BASE COUNT      87 a 102 c 123 g  77 t
ORIGIN

Query Match      11.8%; Score 227; DB 8; Length 389;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 60
|||||
Qy 1690 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 1749
|||||

Db 61 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 120
|||||
Qy 1750 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 1809
|||||

Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 180
|||||
Qy 1810 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 1869
|||||

Db 181 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 231
|||||
Qy 1870 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 1920
|||||

RESULT 3
LOCUS      AA754459      252 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION  AA754459
NID        92801165
KEYWORDS   EST.
SOURCE     rice.
ORGANISM   Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
REFERENCE  1 (bases 1 to 252)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)
COMMENT   Contact: Eun M.Y.
          Department of Cytogenetics
          National Inst. of Agri. Sci. and Tech, RDA
          Suwon, Kyungido, Korea
          Tel: 82 331 290 0301
          Fax: 82 331 290 0307
          Email: myeun@sun20.asti.re.kr
          Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
          University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
          Seq primer: M13 Reverse Primer.
          Unpublished (1998)

FEATURES
source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"

BASE COUNT      87 a 102 c 123 g  77 t
ORIGIN

Query Match      11.8%; Score 227; DB 8; Length 389;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 60
|||||
Qy 1690 GCCAAGTTCCTGCTGCTGATGAGTGTGTACGTCTGCGAGCTGCTCAGCTCTTCTTT 1749
|||||

Db 61 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 120
|||||
Qy 1750 TATGTCAGGAGACACACCTTTCAAAAGACACAGGCTCTTTTCTACCGGAAGTGTCTGG 1809
|||||

Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 180
|||||
Qy 1810 AGCAAGTTGCAAGCATTTGGAATCAGACAGCATTGAAGAGGGTGCAGCTCGGGAGCTG 1869
|||||

Db 181 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 231
|||||
Qy 1870 TCGGAAGCAGAGGTGAGCAGCATCGGAAGCCAGGCGCCGCTGCTGACG 1920
|||||

[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:712562"
/clone_lib="NCL-CPAP-GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
<1..>389
BASE COUNT      87 a 102 c 123 g  77 t
ORIGIN

Query Match      3.0%; Score 58; DB 12; Length 252;
Best Local Similarity 12.8%; Pred. No. 1.27e-46;
Matches 23; Conservative 88; Mismatches 69; Indels 0; Gaps 0;

Db 24 BCGNBVWVCVASHGNVMSVHNCITBGTCHDCCKNVNWSMTWTGTYNWNBSYGDWHYWBVB 83
|||||
Cp 893 GCAGGTGACACACACAGAAACACCGCTCACTCGGTCCACCGCTCCGCCGGTGGGCC 834
|||||

Db 84 NTKVDGNHTRCWRBVRMAHYHDYTCNBYYNNNDYHMHBBMYBTGCMCTCTMWCW 143
|||||
Cp 833 CAGGACCCCTGCCACAGCGCGCTCGGCTCGGCTCAGGGCAGCGCCACGCTGGGCTC 774
|||||

Db 144 HYNKTFASGHHSTNDVKSSTNTWCVTSYDKSMHGYWCSBBVKYHTKVTSTRATRSY 203
|||||
Cp 773 TTGGCAACGCGACACTTCGGCTGGCACTGCCCGCGCGCTCCTCGCACCCGGGCTGGC 714
|||||

RESULT 3
LOCUS      AA754459      252 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION  AA754459
NID        92801165
KEYWORDS   EST.
SOURCE     rice.
ORGANISM   Oryza sativa
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Oryza.
REFERENCE  1 (bases 1 to 252)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)
COMMENT   Contact: Eun M.Y.
          Department of Cytogenetics
          National Inst. of Agri. Sci. and Tech, RDA
          Suwon, Kyungido, Korea
          Tel: 82 331 290 0301
          Fax: 82 331 290 0307
          Email: myeun@sun20.asti.re.kr
          Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
          University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
          Seq primer: M13 Reverse Primer.
          Unpublished (1998)

FEATURES
source
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/notes="Vector: pBluescript SK(+); Site.1: EcoRI; Site.2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"

BASE COUNT      5 a  21 c  12 g  35 t  179 others
ORIGIN

Query Match      2.8%; Score 54; DB 12; Length 252;
Best Local Similarity 8.7%; Pred. No. 7.79e-41;
Matches 18; Conservative 105; Mismatches 85; Indels 0; Gaps 0;

```



[illegible]

Db	401	DCBFTGGDTSBHVCMBCBANADGATBACGKGVGBGASTCTMTHYNCCDCKTCGGSAGV	460
Qy	227	GCCTGGTGTGCGTGGCCTGGAGCGACGGCGGCCCGCCCTCTCTCGCCACGG	286
Db	461	TVNHHDWSMAGGARAGCT	479
Qy	287	TGTCCTCGCTGAAGGAGCT	305
RESULT	8		
LOCUS	AF038250	1287 bp	EST
DEFINITION	Homo sapiens clone ntcon9 mRNA, partial sequence, mRNA sequence.		29-JAN-1998
ACCESSION	AF038250		
NID	92815880		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
	Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 1287)		
AUTHORS	Tripodis, N. and Ragoussis, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-DEC-1997) Division of Medical and Molecular Genetics,		
	Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK		
FEATURES	Location/Qualifiers		
source	1. .1287		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="6"		
	/map="6p21.3"		
	/clone="ntcon9"		
	/dev_stage="fetal"		
	/tissue_type="brain; liver"		
	/note="similar to HSRP20"		
BASE COUNT	349 a 219 c 293 g 361 t	65 others	
ORIGIN			
	Query Match	1.5%;	Score 28; DB 12; Length 1287;
	Best Local Similarity	23.8%;	Pred. No. 5.22e-07;
	Matches	15; Conservative	29; Mismatches 19; Indels 0; Gaps 0;
Db	407	GDTBSHBVCMBCBANADGATBACGKGVGBGASTCTMTHYNCCDCKTCGGSAGVTVNHHD	466
Cp	1426	GGCGAGCAGTGCACCGACGCGCGGGGTCTGTCTCTCTCTCGGGGGCCGCCACAG	1367
Db	467	WSM	469
Cp	1366	AGC	1364
RESULT	9		
LOCUS	AQ008007	317 bp	DNA
DEFINITION	CIT-HSP-2288C3.TF CIT-HSP Homo sapiens genomic clone 2288C3,	GSS	26-JUN-1998
	genomic survey sequence.		
ACCESSION	AQ008007		
NID	93085233		
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
	Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 317)		
AUTHORS	Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,		
	Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,		
	Simon, M. and Venter, J.C.		
TITLE	Use of a random BAC end Sequence Database for Sequence-Ready Map		
JOURNAL	Building		
	Unpublished (1997)		

Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamst@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page:

[http://www.tigr.org/tldb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html)

Seq primer: MJ3-21;

Class: BAC ends.

#### FEATURES

source

Location/Qualifiers

1..317

/organism="Homo sapiens"

/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2:

HindIII"

/db\_xref="taxon:9606"

/clone\_lib="CIT-HSP"

/sex="Male"

/cell\_type="Sperm"

BASE COUNT

59 a 92 c 91 g 75 t

ORIGIN

Query Match

Best Local Similarity 1.4%; Score 27; DB 27; Length 317;

Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db

231 CCTTCCCGAGGTCTCTGCTGAAGGAGTGGGGC 267

QY

275 CCTTCCCGAGGTCTCTGCTGAAGGAGTGGTGGC 311

RESULT 10

LOCUS

H44707 400 bp mRNA EST 31-JUL-1995

DEFINITION

YP24C08.r1 Homo sapiens cDNA clone 188366 5'.

ACCESSION

H44707

NID

9220759

KEYWORDS

EST.

SOURCE  
human clone-188366 library-Soares breast 3NHBst vector-pr7T3D  
(Pharmacia) with a modified polylinker host-DH10B (ampicillin  
resistant) primer-W13RPI Rsite1-Not I Rsite2-Eco RI Adult human.  
1st strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAGTGGAGCGGCCCTTTTCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),  
digested with Not I and cloned into the Not I and Eco RI sites of a  
modified p7T3 vector (Pharmacia). Library went through one round  
of normalization to a Cot = 20. Library constructed by Bento Soares  
and M.Fatima Bonaudo.

ORGANISM

Homo sapiens  
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 400)

REFERENCE

AUTHORS

Hallier,B., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilkon RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 363  
Source: IMAGE Consortium LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Location/Qualifiers

FEATURES

Unpublished

source

1..400

/organism="Homo sapiens"

/clone="188366"

BASE COUNT 75 a 123 c 113 g 79 t 10 others

ORIGIN

Query Match

Best Local Similarity 1.4%; Score 26; DB 16; Length 400;

Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db

28 TNCCTGCCCCGGCCACGTTTGTGAGCAGCTGTAGACCCGAGAGCTCTCGGCTG 80

QY

128 TGTCTGCGCTGGCCACGTTCTGTCGGCGCTGGGGCCCCCAGGCTGGGCTG 180

RESULT 11

LOCUS

AF034177 660 bp mRNA EST 22-DEC-1997

DEFINITION

Homo sapiens ntcon6 contig mRNA, partial sequence, mRNA sequence.

ACCESSION

AF034177

NID

92707739

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 660)  
Tripodis,N. and Ragoussis,J.  
Generation of a transcription map in the region immediately  
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal  
boundary

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 660)

AUTHORS

Tripodis,N. and Ragoussis,J.

TITLE

Direct Submission

JOURNAL

Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,  
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK  
Location/Qualifiers

FEATURES

source

1..660

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="6"

/map="6p21.3"

/clone="ntcon6 contig"

/tissue.type="fetal brain; fetal liver; adult muscle"

/note="similar to CutA"

BASE COUNT 162 a 161 c 173 g 109 t 55 others

ORIGIN

Query Match

Best Local Similarity 1.4%; Score 27; DB 11; Length 660;

Matches 13; Conservative 28; Mismatches 14; Indels 0; Gaps 0;

Db

380 CACSWSRCCYRSYGMVHYCYKCDMSCTTSKSRWYKSRWCDCDCAAG 434

Cp

1419 CAGCTGCACGAGCGAGGGGTCTGTCTCTCTCTCGGGGGCCGACAG 1365

RESULT 12

LOCUS

AF034173 2275 bp mRNA EST 22-DEC-1997

DEFINITION

Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.

ACCESSION

AF034173

NID

92707735

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 2275)  
Tripodis,N. and Ragoussis,J.  
Generation of a transcription map in the region immediately  
centromeric to human MHC across the 6p21.2-6p21.3 chromosomal  
boundary  
Unpublished





**THIS PAGE BLANK (USPTO)**



(TM)

Result No.	Score	Query	ID	Description	Pred. No.	
1	45	2.3	91 9	Q51746	Oligonucleotide probe	3.54e-10
2	45	2.3	91 9	Q51746	Oligonucleotide probe	3.54e-10
C 3	43	2.2	204 1	N81164	Base substituted E.co	3.86e-09
4	40	2.1	114 12	Q70466	Generic DNA sequence	1.32e-07
5	39	2.0	114 12	Q70469	Generic DNA sequence	4.21e-07
6	38	2.0	114 12	Q70468	Generic DNA sequence	1.33e-06
C 7	38	2.0	114 12	Q70467	Generic DNA sequence	1.33e-06
C 8	38	2.0	114 12	Q70468	Generic DNA sequence	1.33e-06
C 9	38	2.0	114 12	Q70465	Generic DNA sequence	1.33e-06
10	39	2.0	178 32	T76405	Human endothelin-1 an	4.21e-07
C 11	38	2.0	204 1	N81164	Base substituted E.co	1.33e-06
C 12	36	1.9	114 12	Q70467	Generic DNA sequence	1.31e-05

```

Query Match      2.3%; Score 45; DB 9; Length 91;
Best Local Similarity 7.3%; Pred. No. 3.54e-10;
Matches 4; Conservative 46; Mismatches 5; Indels 0; Gaps 0;

Db      6  cgcgcgssvhsyyvvhvshhshvhhvhhvsvvvhhvhhvhhvhhvsvv 50
||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      390  CGGGGGCCCCCGAGCGCTTACCACGCGTGCGAGTACTGCCCCAACG 444

RESULT      2

```

Best Local Similarity 20.9%; Pred. NO. 3.86e-09;  
Matches 31; Conservative 57; Mismatches 60; Indels 0; Gaps 0;

Db 45 bcyrraggnycggggywccgcgaaycchvgcgymrttthhyrmbnvyrd 104  
:::|:|||||||::|  
QY 829 TCTGGGCCCCACCGGCGAGCAGCGGTGCAGCATGACCTGTTCCTGTGTGTCTCA 888  
:::|:|||||||::|  
Db 105 ynrsdaaaawcyrrsvkyvdcynachdhdyvvbbvnyvhnnncnccbnnhvchnv 164  
:::|:|||||||::|  
QY 889 CTTGCCAGACCGCGGAGAGACCACCTTGTGGAGGGTGCGCTCTGGCACCGCGCAC 948  
:::|:|||||||::|  
Db 165 hbmhnrmwayvrhdarrddvhocvchcc 192  
:::|:|||||||::|  
QY 949 TCCCACCATCGTGGCGCGCAGCAC 976

RESULT 4

ID	Q70466 standard; DNA; 114 BP.
DT	Q70466; AC
DE	05-APR-1995 (first entry)
KW	Generic DNA sequence to generate a random TSAR-9 peptide library.
DE	TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW	effector domain; concatenated heterofunctional protein; linker;
KW	direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
FT	Key
FT	Location/Qualifiers
FT	misc_feature 55..60
FT	/*tag= a
FT	/note= "this sequence represents 'Z'; Z can be a
FT	sequence of 6, 9 or 12 nucleotides (see
FT	comments")
PN	WT0918318-A.
PD	18-AUG-1994.
PF	01-FEB-1994; U00977.
PR	01-FEB-1994; US-013416.
PR	30-DEC-1993; US-176500.
PR	31-JAN-1994; US-189331.
PI	(UYNC-) UNIV NORTH CAROLINA.
PA	Fowlkes DM. Kay BK;
DR	WPI; 94-279739/34.
DR	P-FSDB; R63152.
PT	Identifying proteins or peptide(s) which bind a ligand - by
PT	screening a recombinant vector library expressing fusion proteins
PT	comprising a binding domain and an effector domain
PS	Disclosure; Page 35: 255pp; English
CC	Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC	Synthetic Affinity Reagents) peptides. This generic formula can also be
CC	represented as follows: X(NNB)1(TGC)(NNB)10(TGC)2(NNB)42(NNB)8(TGC)(NNB)
CC	-9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC	that are not specified further. Other generic sequences are shown in
CC	Q70466-68. Other specific peptides generated by these generic sequences
CC	are shown in R63151-54. TSARs are concatenated heterofunctional proteins
CC	or peptides, comprising at least two functional regions - a binding
CC	domain with affinity for a ligand and a second effector peptide portion
CC	that is chemically or biologically active. They may further comprise a
CC	linker peptide between the 2 domains. The oligonucleotides are also
CC	designed so that the expressed peptide contains 2 or 4 cysteine residues
CC	positioned in, or flanking, the unpredicted or variant residues. These
CC	residues confer some degree of conformational rigidity to the peptides.
CC	The TSARs or compns. comprising a TSAR binding domain can be used in
CC	vivo to deliver a chemically or biologically active moiety, eg, metal
CC	ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC	on the cell. They can also replace the function of macromolecules, eg, for
CC	monoclonal or polyclonal antibodies and therefore circumvent the need for
CC	complex methods of hybridoma formation or in vivo antibody production.
CC	The TSARs are easily characterised and have designed activity allowing
CC	direct and rapid detection in a screening process.
CC	Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Watch 2.1%; Score 40; DB 12; Length 114;  
Best Local Similarity 8.9%; Pred. No. 1.32e-07;  
Matches 10; Conservative 32; Mismatches 70; Indels 0; Gaps 0;

Db		1	tgcnnbnbnbnbnbnbnbnbnbnbnbnbtgcnbnbnbnbnbnbnbnbnbnbn	60
			:     :     :	
QY		149	TGCGGCCTTGGGCCCCAGGGCTGCGCGTGTCAGCGGGACCGGGCTTCC	208
Db		61	nbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn	111
QY		209	GCGGGTGTGGCCCCAAGTGCCTGGTGTGCGTCCCTGGGACGACGGCGC	259

RESULT 7  
ID Q70467 standard; DNA; 114 BP.  
AC Q70467;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR petiole library.





```

DE 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents '2'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PE 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 25pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(GCC)(NNB)16. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns, comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
SQ
Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.31e-05;
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
QY CTCCTCCGTCGCGCAGCCACTACCGAGGTGCTGCGCGCTGCGCAGCTTGTGCGGCGCT 158
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 159 GGGGCCCGAGGCTGCGGCTGTGTGAGCGCGGGGACCCGGCGCTTCGC 210
Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.31e-05;
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
QY CTCCTCCGTCGCGCAGCCACTACCGAGGTGCTGCGCGCTGCGCAGCTTGTGCGGCGCT 158
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 159 GGGGCCCGAGGCTGCGGCTGTGTGAGCGCGGGGACCCGGCGCTTCGC 210
RESULT 13
ID Q70470 standard; DNA; 114 BP.
AC Q70470;
DE 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents '2'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PE 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 25pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(GCC)(NNB)16. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns, comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
SQ
Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.31e-05;
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
QY CTCCTCCGTCGCGCAGCCACTACCGAGGTGCTGCGCGCTGCGCAGCTTGTGCGGCGCT 158
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 159 GGGGCCCGAGGCTGCGGCTGTGTGAGCGCGGGGACCCGGCGCTTCGC 210

```

```

FT misc_feature 55..60
FT /*tag= a
FT /*note= "encoded by 2 (see comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PE 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 36; 25pp; English.
CC Q70470 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)4(CAC)(NNB)82(NNB)6(CAC)(NNB)8
CC -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same
CC as Y) that are not specified further. The peptides generated by this and
CC other generic sequences (Q70471-73) have invariant histidine residues
CC incorporated into variant sequences. TSARs are concatenated
CC heterofunctional proteins or peptides, comprising at least two functional
CC regions - a binding domain with affinity for a ligand and a second
CC effector peptide portion that is chemically or biologically active. They
CC may further comprise a linker peptide between the 2 domains. The TSARs
CC or compsns, comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
CC Sequence 114 BP; 5 A; 10 C; 0 G; 0 T;
SQ
Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 9.8%; Pred. No. 1.31e-05;
Matches 11; Conservative 30; Mismatches 71; Indels 0; Gaps 0;
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
QY 492 CGCGTCGGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 552 GGCTCCCGAGGCTGCGGCTGTGTGAGCGCGGGGACCCGGCGCTTCGC 603
RESULT 14
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DE 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents '2'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PE 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.

```

```

PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
PT P-PSDB: R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)12(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or comps. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
CC Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.31e-05;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 52
Qy 69 GCGCGCGCTCCCGCTCGCGAGCGTGGCTGCTGTCGCGACGCCACTACCGCGAGGT 128
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
Qy 129 GCTGCGCGTGGCCACGTCGTCGCGCGCTGGGGCCCGCAGGCTGGCGGCTG 180

RESULT 15
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PD W09418318-A.
PN 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide

```

```

CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)62(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or comps. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
CC Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 36; DB 12; Length 114;
Best Local Similarity 7.9%; Pred. No. 1.31e-05;
Matches 9; Conservative 31; Mismatches 74; Indels 0; Gaps 0;

Db 1 tgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 60
Cp 1095 TGCYCTTGTGCTCGCTGAGGAGTAGAGGAGTGTGTCGCGGTACACCGGGGACAA 1026
Db 61 nbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbgtc 114
Cp 1025 GGCGTGTCCCGAGGACGTGGTGGCCGCGATGTGGATGGGGGGCCCGCGTGTGC 972

Search completed: Thu Dec 24 15:06:48 1998
Job time : 666 secs.

```

**THIS PAGE BLANK (USPTO)**



(TM)

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1920	100.0	2848	22	US-08-951-	Sequence 13, Applicati	0.00e+00
2	1920	100.0	3798	22	US-08-951-	Sequence 19, Applicati	0.00e+00
3	1912	99.6	7029	22	US-08-911-	Sequence 1, Applicatio	0.00e+00
4	1909	99.4	3855	22	US-08-912-	Sequence 4, Applicatio	0.00e+00
5	1909	99.4	3855	22	US-08-911-	Sequence 18, Applicati	0.00e+00
6	1909	99.4	4015	22	US-08-912-	Sequence 1, Applicatio	0.00e+00
7	1909	99.4	4015	24	US-09-052-	Sequence 1, Applicatio	0.00e+00
8	1909	99.4	4015	21	US-08-834-	Sequence 22, Applicati	0.00e+00
9	1906	99.3	4023	23	US-08-926-	Sequence 35, Applicati	0.00e+00
10	1596	83.1	4029	21	US-08-854-	Sequence 17, Applicati	0.00e+00
11	1596	83.1	4029	21	US-08-851-	Sequence 173, Applicat	0.00e+00
12	1497	78.0	4200	22	US-08-912-	Sequence 6, Applicatio	0.00e+00
13	453	23.6	3496	23	US-09-042-	Sequence 1, Applicatio	0.00e+00
14	281	14.6	3346	23	US-09-036-	Sequence 36, Applicati	2.85e+23
15	259	13.5	535	22	US-08-911-	Sequence 52, Applicati	6.37e+21



Sequence 19, Application US/08951733  
Sequence 19, Application US/08951733  
GENERAL INFORMATION:  
APPLICANT: Harrington, Lea A.  
APPLICANT: Robinson, Murray O.  
TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: One Amgen Center Drive  
CITY: Thousand Oaks  
STATE: CA  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951.733  
FILING DATE: 16-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/873,039  
FILING DATE: 11-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/751,189  
FILING DATE: 15-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Oleski, Nancy A.  
REGISTRATION NUMBER: 34,688  
REFERENCE/DOCKET NUMBER: A-433B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (805) 447-6504  
TELEFAX: (805) 499-8011  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3798 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
SEQUENCE 3798 BP: 613 A; 1310 C; 1213 G; 662 T; 0 OTHER.

Query Match 100.0%; Score 1920; DB 22; Length 3798;  
Best Local Similarity 100.0%; Pred No. 0.00e+00;  
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 CACGGCTCCGGGACGCTGCTGCTGCGCAGCTGGGAAGCCCTGGCCCCGGCCACC 61  
Qy 1 CACGGCTCCGGGACGCTGCTGCTGCGCAGCTGGGAAGCCCTGGCCCCGGCCACC 60  
Db 62 CCGGGATGCGGGGCTCCCGCTGCGGAGCGTGGCGCTCCCTGCTGCGGAGCCACTAC 121  
Qy 61 CCGGGATGCGGGGCTCCCGCTGCGGAGCGTGGCGCTCCCTGCTGCGGAGCCACTAC 120  
Db 122 CGCAGGTGCTGCGCTGCGCAGCTGCTGCGGCGCTGGGCGCCCGAGGCTGGCGCTG 181  
Qy 121 CGCAGGTGCTGCGCTGCGCAGCTGCTGCGGCGCTGGGCGCCCGAGGCTGGCGCTG 180  
Db 182 GTGACGCGGGGACCGCGGCTTCCGCGCGCTGGTGGCCAGTGGCTGGTGGCTG 241  
Qy 181 GTGACGCGGGGACCGCGGCTTCCGCGCGCTGGTGGCCAGTGGCTGGTGGCTG 240  
Db 242 CCCTGGGAGCAGCGCGCGCCCGCCCGCTCCCTTCCTTCGCGCAGGTGCTGCTGCTGAAG 301  
Qy 241 CCCTGGGAGCAGCGCGCGCCCGCCCGCTCCCTTCCTTCGCGCAGGTGCTGCTGCTGAAG 300  
Db 302 GAGCTGGTGGCCGAGTCTGTCAGAGGCTGTCGAGCGCGCGCGGAGAAAGTCTGCTGCC 361  
Qy 301 GAGCTGGTGGCCGAGTCTGTCAGAGGCTGTCGAGCGCGCGCGGAGAAAGTCTGCTGCC 360

Db 362 TTGGCTTTCGGCTGCTGAGACGGGGCCCCGGGGGGCCCCCCCCGAGGCTTTCACACACAGC 421  
Qy 361 TTGGCTTTCGGCTGCTGAGACGGGGCCCCGGGGGGCCCCCCCCGAGGCTTTCACACACAGC 420  
Db 422 GTGGCAGCTACCTGCTCCCAACAGGTACACAGCTACGCGGAGCGGGGAGCGGGGCTGGGG 481  
Qy 421 GTGGCAGCTACCTGCTCCCAACAGGTACACAGCTACGCGGAGCGGGGAGCGGGGCTGGGG 480  
Db 482 CTGCTGCTGCGCGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541  
Qy 481 CTGCTGCTGCGCGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 542 TTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601  
Qy 541 TTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 602 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
Qy 601 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Db 662 TGCAGACGGGCTGGAACCATAGCTCAGGAGGCGGGGCTCCCTGCGGCTGCCAGCC 721  
Qy 661 TGCAGACGGGCTGGAACCATAGCTCAGGAGGCGGGGCTCCCTGCGGCTGCCAGCC 720  
Db 722 CCGGGTGCAGAGGCGCGGGGCGAGTCCAGCGGAAAGTCTGCGGTTGCCAAAGAGGCC 781  
Qy 721 CCGGGTGCAGAGGCGCGGGGCGAGTCCAGCGGAAAGTCTGCGGTTGCCAAAGAGGCC 780  
Db 782 AGCGTGGCGCTGCGCTGAGCGGAGCGGAGCGGCGGCTGCGGAGGCGTCTGGGCGCCAC 841  
Qy 781 AGCGTGGCGCTGCGCTGAGCGGAGCGGAGCGGCGGCTGCGGAGGCGTCTGGGCGCCAC 840  
Db 842 CCGGGCAGGAGCGCTGGACGAGTACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
Qy 841 CCGGGCAGGAGCGCTGGACGAGTACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 902 GCCGAAGAAGCCACTCTTTGGAGGTCGCTCTCTGCGACGCGCCACTCCACCCATCC 961  
Qy 901 GCCGAAGAAGCCACTCTTTGGAGGTCGCTCTCTGCGACGCGCCACTCCACCCATCC 960  
Db 962 GTGGGCGCCAGCAGCAGCGGGGCCCCCATCCACATCGGGGCGCCACGCTCCCTGGGAC 1021  
Qy 961 GTGGGCGCCAGCAGCAGCGGGGCCCCCATCCACATCGGGGCGCCACGCTCCCTGGGAC 1020  
Db 1022 ACGCCTTGTCCCGGTGTACGCGGAGACCAAGCAGCTTCTCTACTCTCAGGCGACAAG 1081  
Qy 1021 ACGCCTTGTCCCGGTGTACGCGGAGACCAAGCAGCTTCTCTACTCTCAGGCGACAAG 1080  
Db 1082 GAGCAGTGGCGGCTCTCTCTACTCTGAGGCGCCAGCCTGAGTGGCGCTCGG 1141  
Qy 1081 GAGCAGTGGCGGCTCTCTCTACTCTGAGGCGCCAGCCTGAGTGGCGCTCGG 1140  
Db 1142 AGGCTGTGAGAGCACTTTCTGGGTTTCAGGCGCTTGGATGCCAGGACTTCCCGCAGG 1201  
Qy 1141 AGGCTGTGAGAGCACTTTCTGGGTTTCAGGCGCTTGGATGCCAGGAGTCCCGCAGG 1200  
Db 1202 TTGCCCCGCTGCGCCAGCGCTACTGGCAATGGGCGCCCTGTTTCTGAGAGCTCTTGGG 1261  
Qy 1201 TTGCCCCGCTGCGCCAGCGCTACTGGCAATGGGCGCCCTGTTTCTGAGAGCTCTTGGG 1260  
Db 1262 AACACACGCGAGTGCCTTACGGGCTCTCTCTAAGACGCACTGCCCGCTCGAGCTGCG 1321  
Qy 1261 AACACACGCGAGTGCCTTACGGGCTCTCTCTAAGACGCACTGCCCGCTCGAGCTGCG 1320  
Db 1322 GTACACCCAGCAGCGGCTGCTGTGCGCGGAGAAAGCCAGGAGTCTGTGGCGGCGCCCC 1381  
Qy 1321 GTACACCCAGCAGCGGCTGCTGTGCGCGGAGAAAGCCAGGAGTCTGTGGCGGCGCCCC 1380  
Db 1382 GAGGAGGAGACAGACACCCCGCTGCGCTGGTGCAGCTGCTCCGCCAGCAGCAGCGCCC 1441  
Qy 1381 GAGGAGGAGACAGACACCCCGCTGCGCTGGTGCAGCTGCTCCGCCAGCAGCAGCGCCC 1440

Db	1442	TGCGAGGTGACGGCTTCCTGCGGGCTGCCTTCGCCCGCTGGTGCCTCCACAGGCCTCTGG	1501
QY	1441	TGCGAGGTGACGGCTTCCTGCGGGCTGCCTTCGCCCGCTGGTGCCTCCACAGGCCTCTGG	1500
Db	1502	GGCTCAGGCACAAAGACCGCTTCCTCAGGAACACCAAGAAGTTTCACTCCCTCGGG	1561
QY	1501	GGCTCAGGCACAAAGACCGCTTCCTCAGGAACACCAAGAAGTTTCACTCCCTCGGG	1560
db	1562	AAGCATGCCAAGCTCTCGTTCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGCT	1621
QY	1561	AAGCATGCCAAGCTCTCGTTCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGCT	1620
Db	1622	TGGCTCGGAGGACCCAGGGGTTGGCTGTGTTCCGGCCCGAGAGCACCCTGTCGCTGAG	1681
QY	1621	TGGCTCGGAGGACCCAGGGGTTGGCTGTGTTCCGGCCCGAGAGCACCCTGTCGCTGAG	1680
Db	1682	GAGATCCTGGCCAAAGTTCTCGCACTGGGTGATGAGTGTACGTCTGAGCTGCTCAGG	1741
QY	1681	GAGATCCTGGCCAAAGTTCTCGCACTGGGTGATGAGTGTACGTCTGAGCTGCTCAGG	1740
Db	1742	TCCTTCTTTTATGTACGGAGACACAGTTTCAAAGAACAGGCTCTTTTCTACCGGAAG	1801
QY	1741	TCCTTCTTTTATGTACGGAGACACAGTTTCAAAGAACAGGCTCTTTTCTACCGGAAG	1800
Db	1802	AGTGTCTGGAGCAAGTTGCAAAAGCATTGAATCAGACAGCACTTGAAGAGGGTGCAGCTG	1861
QY	1801	AGTGTCTGGAGCAAGTTGCAAAAGCATTGAATCAGACAGCACTTGAAGAGGGTGCAGCTG	1860
Db	1862	CGGGAGCTGTCCGAAGCAGAGGTACAGGAGCATCGGGAAGCCAGGCCCGCCTGCTGACG	1921
QY	1861	CGGGAGCTGTCCGAAGCAGAGGTACAGGAGCATCGGGAAGCCAGGCCCGCCTGCTGACG	1920

RESULT 3  
ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.  
AC xxxxxx

```

Sequence 1, Application US/08911312
Sequence 1, Application US/08911312
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William
TITLE OF INVENTION: Telomerase Reverse Transcriptase
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,312
FILING DATE: 14-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017

```

```

CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-00250005
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7029 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 782..4177
CC OTHER INFORMATION: /product= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTERT)"
CC OTHER INFORMATION: /note= "cdna contained in plasmid
CC OTHER INFORMATION: pGRN121"
CC SEQUENCE 7029 bp: 1416 A; 2122 C; 2051 G; 1440 T; 0 OTHER
SQ

```

Query Match	99.6%	Score 1912;	DB 22;	Length 7029;
Best Local Similarity	99.8%;	Pred. No. 0.00e+00;		
Matches 1915;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Db	718	CGCGTCGACGACGGCTGCGTCCGTGCGACACTGGAGAGCCCTGGCCCGGCACCC	777	
Qy	3	CGCGTCGGGACGCGCTGCGTCCGTGCGACACTGGAGAGCCCTGGCCCGGCACCC	62	
Db	778	CGCGATGCGGCGGCTCCCGCTGCCGAGCGTGCCTGCTGCGCAGCCACTACCG	837	
Qy	63	CGCGATGCGGCGGCTCCCGCTGCCGAGCGTGCCTGCTGCGCAGCCACTACCG	122	
Db	838	CGAGGTGCTCGCGTGCACAGTTCTGTCGCGCCTGGGCCCCAGGCTGGCGCTGGT	897	
Qy	123	CGAGGTGCTCGCGTGCACAGTTCTGTCGCGCCTGGGCCCCAGGCTGGCGCTGGT	182	
Db	898	GCACGGGGGACCCGGGGGCTTTCGCGCGCTGGTGGCCAGTGCCTGTTGTCGTGCC	957	
Qy	183	GCACGGGGGACCCGGGGGCTTTCGCGCGCTGGTGGCCAGTGCCTGTTGTCGTGCC	242	
Db	958	CTGGAGCGACGGCGCCCGCCCGCTCTCTCCGCCAGTGTCTGCCTGAAGGA	1017	
Qy	243	CTGGAGCGACGGCGCCCGCCCGCTCTCTCCGCCAGTGTCTGCCTGAAGGA	302	
Db	1018	GCTGGTGCACGAGTGTGTGAGAGGCTGTGCGACGGCGGCGAAGACGTGCTGGGCTT	1077	
Qy	303	GCTGGTGCACGAGTGTGTGAGAGGCTGTGCGACGGCGGCGAAGACGTGCTGGGCTT	362	
Db	1078	CGGCTTCGCGTGTGACAGGGGCGCGGGGGCCCCCGGAGGCTTTCACCAACGCGT	1137	
Qy	363	CGGCTTCGCGTGTGACAGGGGCGCGGGGGCCCCCGGAGGCTTTCACCAACGCGT	422	
Db	1138	CGCAGCTACTGCCCAACACCGTAGCCGACCTCGGGGGAGCGGGGCTGGGGCT	1197	
Qy	423	CGCAGCTACTGCCCAACACCGTAGCCGACCTCGGGGGAGCGGGGCTGGGGCT	482	
Db	1198	GCTGCTCGCGCTGGGCGACGACGTGCTGTTCACTGCTGTCGACGCTGCGCGCTT	1257	







Db 1681 CAAGTTCTGCACTGCTGATCAGTGTGTACGTGCTGAGCTGCTCAGGTCTTTCTTTTA 1740  
QY 1692 CAAGTTCTGCACTGCTGATGAGTGTGTACGTGCTGAGCTGCTCAGGTCTTTCTTTTA 1751  
Db 1741 TGTACGGAGACACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAG 1800  
QY 1752 TGTACGGAGACACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAG 1811  
Db 1801 CAAGTTGCAAGCATTTGGAATCAGACGACTTGAAGAGGTGTCAGCTCCGGAGCTGTC 1860  
QY 1812 CAAGTTGCAAGCATTTGGAATCAGACGACTTGAAGAGGTGTCAGCTCCGGAGCTGTC 1871  
Db 1861 GGAAGCAGAGGTGAGCAGCATCGGAGCCAGGCGCGCCCTGCTGACG 1909  
QY 1872 GGAAGCAGAGGTGAGCAGCATCGGAGCCAGGCGCGCCCTGCTGACG 1920

RESULT 5  
ID US-08-911-312-18 STANDARD; DNA; UNC; 3855 BP.  
AC xxxxx  
DT Sequence 18, Application US/08911312  
CC Sequence 18, Application US/08911312  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase.  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911,312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Elnhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440

CC REFERENCE/DOCKET NUMBER: 015389-002500US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 18:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3855 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56...2476  
CC OTHER INFORMATION: /product= "hprt"  
CC OTHER INFORMATION: /note= "clone #712562"  
CC SEQUENCE 3855 BP: 651 A; 1300 C; 1226 G; 678 T; 0 OTHER.  
  
Query Match 99.4%; Score 1909; DB 22; Length 3855;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 CGACGGCTGCGTCTGCTGCGCACGTGGAAAGCCCTGGCCCGGCGCACCCCGCGCATGCC 60  
QY 12 CGACGGCTGCGTCTGCTGCGCACGTGGAAAGCCCTGGCCCGGCGCACCCCGCGCATGCC 71  
Db 61 GCGCGCTCCCGCTGCGGAGCCGCTGCTTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 120  
QY 72 GCGCGCTCCCGCTGCGGAGCCGCTGCTTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 131  
Db 121 GCGCGTGGCCAGCTGCTGCGGCGCTGGGCGCCAGGGCTGGCGGCTGGTGACGCGGG 180  
QY 132 GCGCGTGGCCAGCTGCTGCGGCGCTGGGCGCCAGGGCTGGCGGCTGGTGACGCGGG 191  
Db 181 GGACCCGCGGCTTCCGCGCGCTGGTGGCCAGTGCCTGCTGCTGCTGCGTGGGAGCG 240  
QY 192 GGACCCGCGGCTTCCGCGCGCTGGTGGCCAGTGCCTGCTGCTGCTGCGTGGGAGCG 251  
Db 241 AGCGCCGCGCCCGCGCGCCCTCTCTCCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCG 300  
QY 252 AGCGCCGCGCCCGCGCGCCCTCTCTCCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCG 311  
Db 301 CCGAGTGTCTGAGAGGCTGTGGAGCGCGCGCGCGGAGAACGTGCTGGGCTTGGCTTGGC 360  
QY 312 CCGAGTGTCTGAGAGGCTGTGGAGCGCGCGCGCGGAGAACGTGCTGGGCTTGGCTTGGC 371  
Db 361 GCTGCTGACGGGCGCGCGGGGCGCCCGAGGCGCTTCCACCGAGGCTTCCACCGAGGCT 420  
QY 372 GCTGCTGACGGGCGCGCGGGGCGCCCGAGGCGCTTCCACCGAGGCTTCCACCGAGGCT 431  
Db 421 CCGTCCCAACAGGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGGCTGCTGCTGCG 480  
QY 432 CCGTCCCAACAGGTGACCGACGCTGCGGGGAGCGGGGCGTGGGGGCTGCTGCTGCG 491  
Db 481 CCGCGTGGGCGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 492 CCGCGTGGGCGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551  
Db 541 GGTCTCCAGCTGCGGCTTACCGAGGTGCGGGCGCGCGCTGTACCGAGCTGCGGCGCTGCCAC 600  
QY 552 GGTCTCCAGCTGCGGCTTACCGAGGTGCGGGCGCGCGCTGTACCGAGCTGCGGCGCTGCCAC 611  
Db 601 TCAGGCGCGGCG 660  
QY 612 TCAGGCGCGGCG 671  
Db 661 CTGGAACCATAGCTGACGAGGCGCGGGTCCCTCTGCGGCTTCCACCGAGGCTGCGGAGCG 720  
QY 672 CTGGAACCATAGCTGACGAGGCGCGGGTCCCTCTGCGGCTTCCACCGAGGCTGCGGAGCG 731  
Db 721 GAGGCGCGGGGCGAGTGCACCGCGGAAGTCTGCCGTTGCCCAAGAGGCCCGCGGCTGGCG 780  
QY 732 GAGGCGCGGGGCGAGTGCACCGCGGAAGTCTGCCGTTGCCCAAGAGGCCCGCGGCTGGCG 791





CC	NAME/KEY:	CDS	Query Match	99.4%	Score	1909;	DB	22;	Length	4015;
CC	LOCATION:	56..3454	Best Local Similarity	100.0%	Pred. No.	0.00e+00;				
CC	OTHER INFORMATION:	/product= "htrr"	Matches	1909;	Conservative	0;	Mismatches	0;	Indels	0;
CC	OTHER INFORMATION:	/note= "human telomerase reverse transcriptase (htrr) catalytic protein component"								
CC	OTHER INFORMATION:	component"								
CC	SEQUENCE	4015 BP; 663 A; 1363 G; 714 T; 0 OTHER.								
DB	1	GCAGCGCTGCGTCTGCTGCGCAGCTGGGAAGCCCTGGCCCCCGCCACACCCCGCGATGCC	60							
QY	12	GCAAGCGTGGCTCTGCTGCGCAGCTGGGAAGCCCTGGCCCCCGCCACACCCCGCGATGCC	71							
DB	61	GCAGCGTCCCGCTGCGGACCGCTGCGCTCCCTGCTGCGCAGCCTACCGCGCAGGTGCT	120							
QY	72	GCGGCTTCCCGCTGCGGACCGCTGCGCTCCCTGCTGCGCAGCCTACCGCGAGGTGCT	131							
DB	121	GCGGCTGGCCACGTTGCTGCGGCGGCTTGGGCGCCCGCAGGCTGGCGGTGCTGTCAGCGCGG	180							
QY	132	GCGGCTGGCCACGTTGCTGCGGCGGCTTGGGCGCCCGCAGGCTGGCGGTGCTGTCAGCGCGG	191							
DB	181	GGACCGCGGCGCTTTCCGCGCGCTGCTGGGCGCCAGTGCCTGGTGTGGTGCCTTGGGACGC	240							
QY	192	GGACCGCGGCGCTTTCCGCGCGCTGCTGGGCGCCAGTGCCTGGTGTGGTGCCTTGGGACGC	251							
DB	241	ACGGCGCGCCCCCGCGCGCCCCCTTCCTTCGCGCAGGTGCTGCTGCTGAAGAGCTGCTGGC	300							
QY	252	ACGGCGCGCCCCCGCGCGCCCCCTTCCTTCGCGCAGGTGCTGCTGCTGAAGAGCTGCTGGC	311							
DB	301	CCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGGAAGACGTGCTGGCTTCGGCTTCGCTTCGC	360							
QY	312	CCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGGAAGACGTGCTGGCTTCGCTTCGCTTCGC	371							
DB	361	GCTGCTGGAGCGGCGCGCGCGGCGCGCGCGGAGGCTTCCACACAGCTGCGCAGCTA	420							
QY	372	GCTGCTGGAGCGGCGCGCGCGGCGCGCGCGGAGGCTTCCACACAGCTGCGCAGCTA	431							
DB	421	CTGCCCCAACACGTGACACGACACTGCGGGGAGCGGGGCTGGTGGTGGTGGTGGTGGTGG	480							
QY	432	CTGCCCCAACACGTGACACGACACTGCGGGGAGCGGGGCTGGTGGTGGTGGTGGTGGTGG	491							
DB	481	CCGCGTGGGCGAGCAGCTGTGTTTACCTGCTGGCAGCTGCGCGCTCTTTGCTGCTGT	540							
QY	492	CCGCGTGGGCGAGCAGCTGTGTTTACCTGCTGGCAGCTGCGCGCTCTTTGCTGCTGT	551							
DB	541	GGTCTCCAGTGGCGCTTACAGAGTGTGGGCGCGCGCTGTACAGCTGCGCGCTGCCAC	600							
QY	552	GGCTCCAGTGGCGCTTACAGAGTGTGGGCGCGCGCTGTACAGCTGCGCGCTGCCAC	611							
DB	601	TCAGGCG	660							
QY	612	TCAGGCG	671							
DB	661	CTGGAACCATAGCTCAGGAGCGCGGCTTCCCTTGGGCTTCCAGCGCGCGCGCGCGCGCGCG	720							
QY	672	CTGGAACCATAGCTCAGGAGCGCGGCTTCCCTTGGGCTTCCAGCGCGCGCGCGCGCGCGCG	731							
DB	721	GAGCGCGCGGCGAGTGCACGCGAAGTCTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCG	780							
QY	732	GAGCGCGCGGCGAGTGCACGCGAAGTCTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCG	791							
DB	781	TGCCCCGTAGCCGAGCGGAGCGGCTTGGGCGAGGCTTCTGGGCGCGCGCGCGCGCGCGCGCG	840							
QY	792	TGCCCCGTAGCCGAGCGGAGCGGCTTGGGCGAGGCTTCTGGGCGCGCGCGCGCGCGCGCGCG	851							
DB	841	CGGTGGACCGAGTACCGTGGTTTCTGCTGTGGTGTACCTGCGCAGACCGCGCGCGCGCGCGCG	900							
QY	852	CGGTGGACCGAGTACCGTGGTTTCTGCTGTGGTGTACCTGCGCAGACCGCGCGCGCGCGCGCG	911							
DB	901	CACCTCTTTGGAGGGTGCCTCTCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	960							

RESULT	7
ID	US-09-
AC	xxxxxx
DT	
DE	Sequen

Sequence 1, Application US/09052919  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Antisense Compositions for Detecting and  
TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/052,919  
FILING DATE: 31-MAR-1998  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

CC LENGTH: 4015 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56..3454  
CC OTHER INFORMATION: /product= "human telomerase reverse  
CC OTHER INFORMATION: transcriptase (hTERT)"  
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.  
  
Query Match 99.4%; Score 1909; DB 24; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.0e+00;  
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1 GCAGCGCTGGTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 60  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
12 GCAGCGCTGGTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGCCACCCCGCGATGCC 71  
Db 61 GCGCGCTCCCGCTGCGCAGCCGCTCCCTGCTGCGCAGCAGCAGTACCGCGAGTGC 120  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
72 GCGCGCTCCCGCTGCGCAGCCGCTCCCTGCTGCGCAGCAGCAGTACCGCGAGTGC 131  
Db 121 GCGCGTGGCCACGTTGCTGGGGCCCTGGGGCCCGCCAGGGCTGGGGCTGTGACGCGG 180  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
132 GCGCGTGGCCACGTTGCTGGGGCCCTGGGGCCCGCCAGGGCTGGGGCTGTGACGCGG 191  
Db 181 GGACCCGGCGGCTTCCGCGCGCTGGTGGCCAGTGGTGGTGGTGGTGGTGGTGG 240  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
192 GGACCCGGCGGCTTCCGCGCGCTGGTGGCCAGTGGTGGTGGTGGTGGTGGTGG 251  
Db 241 ACGCGCCCGCCCGCCCGCCCTCCCTCCCGCAGGTTCTCTGCTGAAGAGTGGTGGC 300  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
252 ACGCGCCCGCCCGCCCGCCCTCCCTCCCGCAGGTTCTCTGCTGAAGAGTGGTGGC 311  
Db 301 CCGAGTCTGCAGAGGCTGTGCGAGCGCGCGGCGAAGACGTGCTGGCCTTCGCTTCG 360  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
312 CCGAGTCTGCAGAGGCTGTGCGAGCGCGCGGCGAAGACGTGCTGGCCTTCGCTTCG 371  
Db 361 GCTGCTGGAGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
372 GCTGCTGGAGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 431  
Db 421 CCTGCCCAACACAGGTGACGACGCTGCGGGGAGCGGGGCGGCGGCGGCGGCGG 480  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
432 CCGTCCCAACACAGGTGACGACGCTGCGGGGAGCGGGGCGGCGGCGGCGGCGG 491  
Db 481 CCGGCTGGGCGAGCAGCTGCTGCTTACCTGTGCGCAGCTGCGCGCTCTTTGTGCT 540  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
492 CCGGCTGGGCGAGCAGCTGCTGCTTACCTGTGCGCAGCTGCGCGCTCTTTGTGCT 551  
Db 541 GGTCTCCAGCTGCGCCTACCGAGTGTGCGGCGCGCGCTGTACAGCTCGCGCTGC 600  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
552 GGTCTCCAGCTGCGCCTACCGAGTGTGCGGCGCGCGCTGTACAGCTCGCGCTGC 611  
Db 601 TCAGGCGCGGCG 660  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
612 TCAGGCGCGGCG 671  
Db 661 CTGGAACCATAGGCTCAGGAGCGCGGCGTCCCGCTGGGCGCTGCCAGCCCGG 720  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
672 CTGGAACCATAGGCTCAGGAGCGCGGCGTCCCGCTGGGCGCTGCCAGCCCGG 731  
Db 721 GAGCGCGGCGGCGAGTCCAGCGAGCTGTGCGGTTGCCCAAGAGCGCGCGGCG 780  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
732 GAGCGCGGCGGCGAGTCCAGCGAGCTGTGCGGTTGCCCAAGAGCGCGCGGCG 791  
Db 781 TGCCCTTGAGCGGAGCGCGCGCTTGGGCGAGGGGTCTTGGGCGCCACCGCGG 840  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
792 TGCCCTTGAGCGGAGCGCGCGCTTGGGCGAGGGGTCTTGGGCGCCACCGCGG 851  
Db 841 CGTGGACCGAGTGACCGTGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

QY 852 GCGTGGACCGAGTACCGTGGTTCTGTGTGTGTACCTCCAGACCCCGGAGAGGC 911  
Db 901 CACCTCTTTGAGGGTGGCTTCTGTGGACGCGCCACTCCACCCATCCGTGGGCCGCCA 960  
QY 912 CACCTCTTTGAGGGTGGCTTCTGTGGACGCGCCACTCCACCCATCCGTGGGCCGCCA 971  
Db 961 GCACACGCGGGCCCGCCATCCATCCGCGCCACCGAGTCCCTGGGACAGCCTTGTCC 1020  
QY 972 GCACACGCGGGCCCGCCATCCATCCGCGCCACCGAGTCCCTGGGACAGCCTTGTCC 1031  
Db 1021 CCCGGTGTACCCGAGACCAAGCACTTCTCTACTCTCTCAGCGCAACAGGAGCAGCTGCG 1080  
QY 1032 CCCGGTGTACCCGAGACCAAGCACTTCTCTACTCTCTCAGCGCAACAGGAGCAGCTGCG 1091  
Db 1081 GCCCTCTCTCTACTAGCTCTCTGAGGCCAGCCTGACTGGCCCTCGGAGGCTCTGTGA 1140  
QY 1092 GCCCTCTCTCTACTAGCTCTCTGAGGCCAGCCTGACTGGCCCTCGGAGGCTCTGTGA 1151  
Db 1141 GACCATCTTTCTGGTTCAGGCGCCTGGATGCCAGGAGTCCCGCAGGTTGCCCGCCCT 1200  
QY 1152 GACCATCTTTCTGGTTCAGGCGCCTGGATGCCAGGAGTCCCGCAGGTTGCCCGCCCT 1211  
Db 1201 GCCCAGCGCTACTGGCAAAATGCGGCCCTCTGTGTGAGCTGTCTGGGAACACACGCGCA 1260  
QY 1212 GCCCAGCGCTACTGGCAAAATGCGGCCCTCTGTGTGAGCTGTCTGGGAACACACGCGCA 1271  
Db 1261 GTGCCCTCTACGGGTGTCTCTCAAGAGCACTGCCCGCTGCGAGCTGCGGTCAACCCCGAG 1320  
QY 1272 GTGCCCTCTACGGGTGTCTCTCAAGAGCACTGCCCGCTGCGAGCTGCGGTCAACCCCGAG 1331  
Db 1321 AGCCGGTGTCTGTCGCGGGAGAGCCCGAGGCTCTGTGCGCGCCCGCAGGAGGAGGA 1380  
QY 1332 AGCCGGTGTCTGTCGCGGGAGAGCCCGAGGCTCTGTGCGCGCCCGCAGGAGGAGGA 1391  
Db 1381 CACAGACCCCGTCCCTGTGTGAGCTGTCTCCGCGAGCAGACGCCCTGCGAGGTGTA 1440  
QY 1392 CACAGACCCCGTCCCTGTGTGAGCTGTCTCCGCGAGCAGACGCCCTGCGAGGTGTA 1451  
Db 1441 CGGCTTCTGTCGCGGCCCTGCTCGCGGCCCTGTGTGCGGCCCTCTGGGCTCCAGGCA 1500  
QY 1452 CGGCTTCTGTCGCGGCCCTGCTCGCGGCCCTGTGTGCGGCCCTCTGGGCTCCAGGCA 1511  
Db 1501 CAACGACCGCGCTTCTCAGAGACACCAAGAGTTCTATCTCCCTGGGGAAGCATGCCAA 1560  
QY 1512 CAACGACCGCGCTTCTCAGAGACACCAAGAGTTCTATCTCCCTGGGGAAGCATGCCAA 1571  
Db 1561 GCTCTCGCTGAGGAGCTGAGCTGGAAGATGAGCTGCGGAGTCCGCTTGGCTGCCGAG 1620  
QY 1572 GCTCTCGCTGAGGAGCTGAGCTGGAAGATGAGCTGCGGAGTCCGCTTGGCTGCCGAG 1631  
Db 1621 GAGCCCGAGGGTGGCTGTGTTCGCGCCGAGACACCGCTGCGGTGAGGAGATCCCTGGC 1680  
QY 1632 GAGCCCGAGGGTGGCTGTGTTCGCGCCGAGACACCGCTGCGGTGAGGAGATCCCTGGC 1691  
Db 1681 CAAGTTCCTGACCTGGCTGATGATGTGATGATGATGATGATGATGATGATGATGATGAT 1740  
QY 1692 CAAGTTCCTGACCTGGCTGATGATGTGATGATGATGATGATGATGATGATGATGATGAT 1751  
Db 1741 TGTCACGAGACACGCTTCAAAAGACAGCTCTTTTCTACCGGAGAGATGTCGTGAG 1800  
QY 1752 TGTCACGAGACACGCTTCAAAAGACAGCTCTTTTCTACCGGAGAGATGTCGTGAG 1811  
Db 1801 CAAGTTCGAAAGCATTTGAATCAGACAGCTTGAAGAGGTGAGCTGCGGGAGCTGTC 1860  
QY 1812 CAAGTTCGAAAGCATTTGAATCAGACAGCTTGAAGAGGTGAGCTGCGGGAGCTGTC 1871  
Db 1861 GGAAGCAGAGTCAAGCAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909  
QY 1872 GGAAGCAGAGTCAAGCAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920

RESULT 8

US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.  
xxxxxx  
Sequence 224, Application US/08854050  
Sequence 224, Application US/08854050  
GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Novel telomerase  
CC NUMBER OF SEQUENCES: 225  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 224:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4015 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cdna  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56..3454 /product= "hTERT"  
CC OTHER INFORMATION: /note= "human telomerase reverse  
CC OTHER INFORMATION: transcriptase (hTERT) catalytic protein  
CC OTHER INFORMATION: component"  
SQ SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.  
Query Match 99.4%; Score 1909; DB 21; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	1	GCAGCGTGGCTCCTGCTGCGCACGTGGAAAGCCCTTGCCCCCGCGCCACACCCCGCGATGCC	60
Qy	12	GCAGCGTGGCTCCTGCTGCGACGTGGGAAAGCCCTTGCCCCCGCGCCACACCCCGCGATGCC	71
Db	61	GCAGCGTCCCGCTGCGAGCGTGGCTCCTGCTGCGCAGCCACCTACCGCAGAGTGCT	120
Qy	72	GCAGCGTCCCGCTGCGAGCGTGGCTCCTGCTGCGCAGCCACCTACCGCAGAGTGCT	131
Db	121	GCGCGTGGCCACGTTGCTGCGCGGCTTGGGCCCCAGAGGCTTGGCGGCTGGTGAGCGCGG	180
Qy	132	GCGCGTGGCCACGTTGCTGCGCGGCTTGGGCCCCAGAGGCTGGTGAGCGCGG	191
Db	181	GGACCCGGGGCTTCCCGCGGCTGGTGCGCCAGTGCCTGCTGCTGCTGGCGCTGGGACGC	240
Qy	192	GGACCCGGGGCTTCCCGCGGCTGGTGCGCCAGTGCCTGCTGCTGCTGGCGCTGGGACGC	251
Db	241	ACGGCGCCCCCGCGGCCCTCTCTCCGCGAGTGTCCTGAGGAGCTGCTGGC	300
Qy	252	ACGGCGCCCCCGCGGCCCTCTCTCTCCGCGAGTGTCCTGAGGAGCTGCTGGC	311
Db	301	CCGAGTCTGACAGAGCTGTGGAGCGGCGGCGAAGAACCTGCTGCCTTCGGCTTCGC	360
Qy	312	CCGAGTCTGACAGAGCTGTGGAGCGGCGGCGAAGAACCTGCTGCCTTCGGCTTCGC	371
Db	361	GCTGCTGGAGGGGCCCCGGGGGCCCCCGAGGCGCTTACACACAGCGTGGCGAGCTA	420
Qy	372	GCTGCTGGAGCGGGGCCCCGGGGGCCCCCGAGGCGCTTACACACAGCTGGCGAGCTA	431
Db	421	CCTGCCCAACACAGCTCACACAGCACTCGGGGGAGCGGCGCTGGGGCTGCTGCTGGC	480
Qy	432	CCTGCCCAACACAGCTGACCGAGCGACTCGGGGGAGCGGCGCTGGGGCTGCTGCTGGC	491
Db	481	CCGCGTGGGCGACAGCTGTGTTTACACTGCTGCACAGCTGCGCGCTCTTTGTGCTGGT	540
Qy	492	CCGCGTGGGCGACAGCTGTGTTTACACTGCTGCACAGCTGCGCGCTCTTTGTGCTGGT	551
Db	541	GGCTCCAGCTGGCGCTACAGAGTGTGGGGCGCGCGCTGTACAGCTCGGCGCTGCCAC	600
Qy	552	GGCTCCAGCTGGCGCTACAGAGTGTGGGGCGCGCGCTGTACAGCTCGGCGCTGCCAC	611
Db	601	TCAGGCCCCCGCCCGCCACAGCTAGTGGACCCCGAAGGCGTCTGGATGCGCAACGGGC	660
Qy	612	TCAGGCCCCCGCCCGCCACAGCTAGTGGACCCCGAAGGCGTCTGGATGCGCAACGGGC	671
Db	661	CTGGAACCATAGCGTCAGGAGCGCGGCTCCCTCTGGGCTGCCAGCCCGGCTGCGAG	720
Qy	672	CTGGAACCATAGCGTCAGGAGCGCGGCTCCCTCTGGGCTGCCAGCCCGGCTGCGAG	731
Db	721	GAGCGCGGGGCGAGTGCAGCGAGTCTGCCGTTGCCAAGAGGCCCAGGCGTGGCGC	780
Qy	732	GAGCGCGGGGCGAGTGCAGCGAGTCTGCCGTTGCCAAGAGGCCCAGGCGTGGCGC	791
Db	781	TGCCCTTGAGCCGGAGCGACGCCCTTGGGCGAGGGTCTCTGGGCGCCACCCGGGCGAGAC	840
Qy	792	TGCCCTTGAGCCGGAGCGACGCCCTTGGGCGAGGGTCTCTGGGCGCCACCCGGGCGAGAC	851
Db	841	GCSTGGACCGAGTGACCGTGGTTTCTGTGTGTGTACCTGCCAGACCCGCCGAAGAC	900
Qy	852	GCSTGGACCGAGTGACCGTGGTTTCTGTGTGTGTGTACCTGCCAGACCCGCCGAAGAC	911
Db	901	CACCTCTTTGAGGGTGGCTCTGTGGACGCGGCACCTCCAGCCCTCTGGGCGCGCA	960
Qy	912	CACCTCTTTGAGGGTGGCTCTGTGGACGCGGCACCTCCAGCCCTCTGGGCGCGCA	971
Db	961	GCACACCGGGGCCCCCATCCACATCGCGGCCACAGCTCCCTGGGACAGCGCTTGCTCC	1020
Qy	972	GCACACCGGGGCCCCCATCCACATCGCGGCCACAGCTCCCTGGGACAGCGCTTGCTCC	1031
Db	1021	CCGGGTGTACCGAGACACAGCACTTCTCTACTCTCTACGGCGACAGAGCACTGGC	1080
Qy	1032	CCGGGTGTACCGAGACACAGCACTTCTCTACTCTCTACGGCGACAGAGCACTGGC	1091

RESULT	9	
ID	US-09-026-981-35	STANDARD; DNA; UNC; 4023 BP.
AC	AC	xxxxxx
DT		
DE	Sequence 35,	Application US/09026981
CC	Sequence 35,	Application US/09026981
CC	GENERAL INFORMATION:	
CC	APPLICANT:	Counter, Christopher M.
CC	APPLICANT:	Meyerson, Matthew
CC	APPLICANT:	Weinberg, Robert A.
CC	TITLE OF INVENTION:	Telomerase Catalytic Subunit Gene and
CC	NUMBER OF SEQUENCES:	52
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Hamilton, Brook, Smith & Reynolds, P.C.
CC	STREET:	Two Militia Drive
CC	CITY:	Lexington

CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02173  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/026,981  
CC FILING DATE: 20-FEB-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/064,322  
CC FILING DATE: 30-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/055,762  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/054,549  
CC FILING DATE: 01-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/047,151  
CC FILING DATE: 20-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/038,750  
CC FILING DATE: 20-FEB-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Granahan, Patricia  
CC REGISTRATION NUMBER: 32,227  
CC REFERENCE/DOCKET NUMBER: WH197-11p4AM  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 781-861-6240  
CC TELEFAX: 781-861-9540  
CC INFORMATION FOR SEQ ID NO: 35:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4023 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC SEQUENCE 4023 BP: 668 A; 1363 C; 1277 G; 715 T; 0 OTHER.

Query Match 99.3%; Score 1906; DB 23; Length 4023;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1908; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 3 GGCAGCGTGGCTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 62  
|||||  
Qy 11 GGCAGCGTGGCTCTGCTGCGCAGCTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 70  
|||||  
Db 63 GCGCGCTCCCGCTGCGCGAGCGCTGCGCTCCCTGCTGCGCAGCAGCTACCGCGAGTGC 122  
|||||  
Qy 71 GCGCGCTCCCGCTGCGCGAGCGCTGCGCTCCCTGCTGCGCAGCAGCTACCGCGAGTGC 130  
|||||  
Db 123 TGCCGCTGGCCACGTTCTGCGCGCCCTGGGGCCCGCCAGGGCTGGCGGCTGGTGCAGCGG 182  
|||||  
Qy 131 TGCCGCTGGCCACGTTCTGCGCGCCCTGGGGCCCGCCAGGGCTGGCGGCTGGTGCAGCGG 190  
|||||  
Db 183 GGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGTGCTGGTGGCTGGCCCTGGGAGC 242  
|||||  
Qy 191 GGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGTGCTGGTGGCTGGCCCTGGGAGC 250  
|||||  
Db 243 CACGCGCCCGCCCGCCCGCCCTCTTCCGCGAGGTGTCTGCTGAAGAGCTGGTGG 302  
|||||  
Qy 251 CACGCGCCCGCCCGCCCGCCCTCTTCCGCGAGGTGTCTGCTGAAGAGCTGGTGG 310  
|||||  
Db 303 CCGAGTGTGAGAGGTGTGCGAGCGCGCGCGAAGAGCTGTGCGCTTGGCTTCG 362  
|||||  
Qy 311 CCGAGTGTGAGAGGTGTGCGAGCGCGCGCGAAGAGCTGTGCGCTTGGCTTCG 370  
|||||  
Db 363 CGCTGCTGGAGCT 422  
|||||  
Qy 371 CGCTGCTGGAGCT 430  
|||||

Db 423 ACCTGCCCAACACAGGTGACCGACCCACTGCGGGGGAGCGGGGCGCTGGGGCTGTGTTC 482  
|||||  
Qy 431 ACCTGCCCAACACAGGTGACCGACCCACTGCGGGGGAGCGGGGCGCTGGGGCTGTGTTC 490  
|||||  
Db 483 GCGCGTGGCGGAGCGTGTGTTCACCTGCTGGCAGCCTGCGCGCTCTTTTGTGTGG 542  
|||||  
Qy 491 GCGCGTGGCGGAGCGTGTGTTCACCTGCTGGCAGCCTGCGCGCTCTTTTGTGTGG 550  
|||||  
Db 543 TGGCTCCCACTAGCTGCGCTTACCAAGTGTGCGGGCGCGCGCTGTACCAAGTGTGGCGCTGCCA 602  
|||||  
Qy 551 TGGCTCCCACTAGCTGCGCTTACCAAGTGTGCGGGCGCGCGCTGTACCAAGTGTGGCGCTGCCA 610  
|||||  
Db 603 CTAGAGCGCGCGCGCGCGCGCGCTAGTGTGACCCCGAAGCGCTGTGGATGTGCGAAGCGG 662  
|||||  
Qy 611 CTAGAGCGCGCGCGCGCGCGCTAGTGTGACCCCGAAGCGCTGTGGATGTGCGAAGCGG 670  
|||||  
Db 663 CCTGGAACCATAGCTGAGGAGCGCGGCTGCGCGCTGCGGGCTGCCAGCGCGGGTGGGA 722  
|||||  
Qy 671 CCTGGAACCATAGCTGAGGAGCGCGGCTGCGCGCTGCGGGCTGCCAGCGCGGGTGGGA 730  
|||||  
Db 723 GGAGCGCGGGGCGAGTGCACAGCGAAGTGTGCGGTTCGCCAAGAGCGCGCGGCTGGCG 782  
|||||  
Qy 731 GGAGCGCGGGGCGAGTGCACAGCGAAGTGTGCGGTTCGCCAAGAGCGCGCGGCTGGCG 790  
|||||  
Db 783 CTGCCCCGTAGCGCGGAGCGCGCGCTTGGGCGAGGGTCTTGGGCGCGCGCGCGCGCGAGGA 842  
|||||  
Qy 791 CTGCCCCGTAGCGCGGAGCGCGCGCTTGGGCGAGGGTCTTGGGCGCGCGCGCGCGCGAGGA 850  
|||||  
Db 843 GCGCTGAGCGCGAGTGCACGCGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 902  
|||||  
Qy 851 GCGCTGAGCGCGAGTGCACGCGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910  
|||||  
Db 903 CCACCTCTTTGGAGGTTGCGCTCTGTGCGACGCGCGCGCTCTCCACCGCTCTCCCGCGCGCGCC 962  
|||||  
Qy 911 CCACCTCTTTGGAGGTTGCGCTCTGTGCGACGCGCGCGCTCTCCACCGCTCTCCCGCGCGCC 970  
|||||  
Db 963 AGCACCGCGCGCGCGCGCGCGCTTCCACATCGGCGCGCGCGCGCGCGCGCGCGCGCGCTTGT 1022  
|||||  
Qy 971 AGCACCGCGCGCGCGCGCGCGCTTCCACATCGGCGCGCGCGCGCGCGCGCGCGCGCGCTTGT 1030  
|||||  
Db 1023 CCGCGGTGTAGCGCGAGCGCGCGCTTCTCTACTCTCTAGGCGCGCGCGCGCGCGCGCGCGCGCG 1082  
|||||  
Qy 1031 CCGCGGTGTAGCGCGAGCGCGCGCTTCTCTACTCTCTAGGCGCGCGCGCGCGCGCGCGCGCGCG 1090  
|||||  
Db 1083 GCGCGCTCTCTCTACTCTAGGCGCGCGCGCTGACTGCGCGCTGCGAGGCTGTGTGG 1142  
|||||  
Qy 1091 GCGCGCTCTCTCTACTCTAGGCGCGCGCGCTGACTGCGCGCTGCGAGGCTGTGTGG 1150  
|||||  
Db 1143 AGACCATCTTTCTGGGTTCCAGGCGCGCTGGATGCGAGGAGTCTCCCGCGAGTGTGCCCGCGC 1202  
|||||  
Qy 1151 AGACCATCTTTCTGGGTTCCAGGCGCGCTGGATGCGAGGAGTCTCCCGCGAGTGTGCCCGCGC 1210  
|||||  
Db 1203 TGCCCCAGCGCTACTTGGCAATGCGGCGCGCTTCTTCTGGAGCTCTTGGGAGCGCGCGCGCG 1262  
|||||  
Qy 1211 TGCCCCAGCGCTACTTGGCAATGCGGCGCGCTTCTTCTGGAGCTCTTGGGAGCGCGCGCGCG 1270  
|||||  
Db 1263 AGTGGCGCTTACGCGGCTCTCTCAAGAGCGCTGCGCGCTGCGAGCTGCGGTGTACCGCGAG 1322  
|||||  
Qy 1271 AGTGGCGCTTACGCGGCTCTCTCAAGAGCGCTGCGCGCTGCGAGCTGCGGTGTACCGCGAG 1330  
|||||  
Db 1323 CAGCGCGTGTGTGTGCCCGGAGAGCGCGCGCTTGTGGCGCGCGCGCGCGCGCGCGCGAGGAGG 1382  
|||||  
Qy 1331 CAGCGCGTGTGTGTGCCCGGAGAGCGCGCGCTTGTGGCGCGCGCGCGCGCGCGCGCGAGGAGG 1390  
|||||  
Db 1383 ACACAGAGCGCGCGCGCGCTGTGTGAGCTGTCTCCGCGAGCAGCAGCGCGCGCGCGCGAGTGT 1442  
|||||  
Qy 1391 ACACAGAGCGCGCGCGCGCTGTGTGAGCTGTCTCCGCGAGCAGCAGCGCGCGCGCGAGTGT 1450  
|||||  
Db 1443 ACAGCTTGTGCGCGCGCTGCGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1502  
|||||  
Qy 1451 ACAGCTTGTGCGCGCGCTGCGCGCGCGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1510  
|||||







CC LOCATION: 1..4029  
CC OTHER INFORMATION: /note= "preliminary sequence for  
CC OTHER INFORMATION: human TRT cDNA insert of  
CC OTHER INFORMATION: plasmid pGRN121"  
SQ SEQUENCE 4029 BP; 687 A; 1342 C; 1255 G; 720 T; 25 OTHER.

Query Match 83.1%; Score 1596; DB 21; Length 4029;  
Best Local Similarity 95.7%; Pred. No. 0.00e+00;  
Matches 1828; Conservative 0; Mismatches 75; Indels 7; Gaps 6;

Db 1 GCAGCGCTGCGTCTGCGACGTGGAGCCCTGCGCCCGGACCCCGCGCATGCC 60  
Qy 12 GCAGCGCTGCGTCTGCGACGTGGAGCCCTGCGCCCGGACCCCGCGCATGCC 71  
Db 61 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCGTGGCGACCCACTACCGGAGTGCT 120  
Qy 72 GCGCGCTCCCGCTCCGAGCGCTGCGTCTGCGTGGCGACCCACTACCGGAGTGCT 131  
Db 121 GCGCGTGGCGACGTGCTGCGCGGCTGGGGCCCGAGGCTGGCGGTGGTGCAGCGGG 180  
Qy 132 GCGCGTGGCGACGTGCTGCGCGGCTGGGGCCCGAGGCTGGCGGTGGTGCAGCGGG 191  
Db 181 GGACCCGCGGCTTTCGCGCGGCTGGGCCCCANTGCTGCTGCGTGGCGANGN 240  
Qy 192 GGACCCGCGGCTTTCGCGCGGCTGGGCCCCAGTGGCTGGTGGCGTGGCGGCGC 251  
Db 241 ANGGCGCCCGCGCGGCTTTCGCGCGGCTGGGCCCCAGTGGCTGGCGTGGCG 300  
Qy 252 ACGGCGCGCGCGCGGCTTTCGCGCGGCTGGGCCCCAGTGGCTGGCGTGGCG 311  
Db 301 CGGAGTGTGCGANGCTGTGGGCGGCGCGGCTGGCGTGGCGTGGCGTGGCG 360  
Qy 312 CGGAGTGTGCGANGCTGTGGGCGGCGCGGCTGGCGTGGCGTGGCGTGGCG 371  
Db 361 GGTGTGCGAGCGGCGCGGCGGCGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 420  
Qy 372 GGTGTGCGAGCGGCGCGGCGGCGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 431  
Db 421 CTGCGCCGCGCGCGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 480  
Qy 432 CTGCGCCGCGCGCGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 491  
Db 481 CGCGTGGCGAGCGTGTGGTTCACCTGTGCGAGCGTGGCGTGGCGTGGCGTGGCG 540  
Qy 492 CGCGTGGCGAGCGTGTGGTTCACCTGTGCGAGCGTGGCGTGGCGTGGCGTGGCG 551  
Db 541 GGTGTGCGAGCGTGTGGTTCACCTGTGCGAGCGTGGCGTGGCGTGGCGTGGCG 600  
Qy 552 GGTGTGCGAGCGTGTGGTTCACCTGTGCGAGCGTGGCGTGGCGTGGCGTGGCG 611  
Db 601 TCAGGCGCGCGCGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 658  
Qy 612 TCAGGCGCGCGCGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 671  
Db 659 CTGGAACCATAGCGTGGGAGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 717  
Qy 672 CTGGAACCATAGCGTGGGAGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCG 731  
Db 718 GAGGCGCGGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 777  
Qy 732 GAGGCGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 791  
Db 778 TGCCCTCTGAGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 837  
Qy 792 TGCCCTCTGAGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 851  
Db 838 GCGTGGACCGAGTACCGTGGTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 897  
Qy 852 GCGTGGACCGAGTACCGTGGTGTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 911  
Db 898 CACCTCTTGGAGGTGGCTCTGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCG 957  
Qy 912 CACCTCTTGGAGGTGGCTCTGCGGCGGCGGCGGCTGGCGTGGCGTGGCGTGGCG 971

## RESULT 12

ID US-08-912-951-6 STANDARD; DNA; UNC; 4200 BP.

AC xxxxxx

DT

DE Sequence 6, Application US/08912951

CC Sequence 6, Application US/08912951

GENERAL INFORMATION:



CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITLE OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-0026000S  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4200 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 4200 BP; 525 A; 1187 C; 1138 G; 644 T; 706 OTHER.

Query Match 78.0%; Score 1497; DB 22; Length 4200;  
Best Local Similarity 94.0%; Pred. No. 0.00e+00;  
Matches 1630; Conservative 0; Mismatches 0; Indels 104; Gaps 1;

Db 2257 GGCAGCGCTGGCTCTGCTGCGCAGTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 2316  
|||||  
Qy 11 GGCAGCGCTGGCTCTGCTGCGCAGTGGGAAGCCCTGGCCCGGCCACCCCGCGATGC 70  
|||||

Db 2317 GCGCGCTCCCGCTGCGGAGCCGTGGCTCCCTGCTGCGGAGCAGCTACCGCGAGGTGC 2376  
|||||  
Qy 71 GCGCGCTCCCGCTGCGGAGCCGTGGCTCCCTGCTGCTGCGCAGCAGCTACCGCGAGGTGC 130  
|||||

Db 2377 TGGCGCTGGCCACGTTGCTGCGGGGCGCTTGGGGCCCGCCAGAGGCTGGCGGCTGGTGCAGGCGC 2436  
|||||  
Qy 131 TGGCGCTGGCCACGTTGCTGCGGGGCGCTTGGGGCCCGCCAGGAGCTGGCGGCTGGTGCAGGCGC 190  
|||||

Db 2437 GGCACCGCGCGCTTTCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2496  
|||||  
Qy 191 GGCACCGCGCGCTTTCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 250  
|||||

Db 2497 CAGGGCG 2556  
|||||  
Qy 251 CAGGGCG 288  
|||||

Db 2557 GGCTGGGGTTGAGGGGCG 2616  
|||||  
Qy 288 ----- 288

Db 2617 CAGGGCGCTTCCCCCGCAGGTGCTGCTGCGTGAAGGAGCTGCTGGCGCGCGAGTGTGCGAGAG 2676  
|||||  
Qy 289 -----TCTGCTGCTGAAGGAGCTGCTGGCGCGCGAGTGTGCGAGAG 326  
|||||

Db 2677 GCTGTGCGAGCG 2736  
|||||  
Qy 327 GCTGTGCGAGCG 386  
|||||

Db 2737 CCGGGGGGCG 2796  
|||||  
Qy 387 CCGGGGGGCG 446  
|||||

Db 2797 GACCGAGCGACTGCGGGGGGAGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCG 2856  
|||||  
Qy 447 GACCGAGCGACTGCGGGGGGAGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCGGGGGGCG 506  
|||||

Db 2857 CGTGTGCTTCACTGCTGCGCAGCTGCGCGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2916  
|||||  
Qy 507 CGTGTGCTTCACTGCTGCGCAGCTGCGCGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566  
|||||

Db 2917 CTACAGGTGTGCGGGCGCGCGCTGTACCAAGCTGCGGCGCTGCCACTCAGCGCGCGCGCGCGCGCG 2976  
|||||  
Qy 567 CTACAGGTGTGCGGGCGCGCGCTGTACCAAGCTGCGGCGCTGCCACTCAGCGCGCGCGCGCGCGCG 626  
|||||

Db 2977 GGCACAGCTAGTGGAGCG 3036  
|||||  
Qy 627 GGCACAGCTAGTGGAGCG 686  
|||||

Db 3037 CAGGAGCGCGGGTCCCTGCGGCGCTGCCAGCCCGCGGCTGCGAGGAGCGCGGGGCGAG 3096  
|||||  
Qy 687 CAGGAGCGCGGGTCCCTGCGGCGCTGCCAGCCCGCGGCTGCGAGGAGCGCGGGGCGAG 746  
|||||

Db 3097 TGGCAGCGAAGTGTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3156  
|||||  
Qy 747 TGGCAGCGAAGTGTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 806  
|||||

Db 3157 GCGAGCGCGGTGGGCGAGGGGTCTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 3216  
|||||  
Qy 807 GCGAGCGCGGTGGGCGAGGGGTCTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 866  
|||||

Db 3217 CCGTGGTTTCTGTGTGTGTCACTGCGCAGACCGCGCGGAGAGCCACCTTTTGGAGGG 3276  
|||||  
Qy 867 CCGTGGTTTCTGTGTGTGTCACTGCGCAGACCGCGCGGAGAGCCACCTTTTGGAGGG 926  
|||||

Db 3277 TGGGCTCTGTGGCAGCG 3336  
|||||  
Qy 927 TGGGCTCTGTGGCAGCG 986  
|||||

Db 3337 CCATCCACATCGCGGCG 3396  
|||||  
Qy 987 CCATCCACATCGCGGCG 1046  
|||||

Db 3397 GACCAACACTTCTCTACTCTCTAGCGCGCAGAGGAGCAGCTGGGGCGCTTCTTCTTCTTCTTCTTCT 3456  
|||||  
Qy 1047 GACCAACACTTCTCTACTCTCTAGCGCGCAGAGGAGCAGCTGGGGCGCTTCTTCTTCTTCTTCTTCT 1106  
|||||

Db 3457 CAGCTCTCTGAGGCCAGCCTGACTGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGG 3516  
QY 1107 CAGCTCTCTGAGGCCAGCCTGACTGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGG 1166  
Db 3517 TTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCCAGCGCTACTG 3576  
QY 1167 TTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCCAGCGCTACTG 1226  
Db 3577 GCAATATGGGCCCTGTTTCTTGAGCTGCTTGGAGTACACGCGAGTCCCTTACGGGGT 3636  
QY 1227 GCAATATGGGCCCTGTTTCTTGAGCTGCTTGGAGTACACGCGAGTCCCTTACGGGGT 1286  
Db 3637 GCTCTCAAGACGCTGCTGCGCTCGAGCTGCGTTCACCCAGCAGCGGTGCTGTGC 3696  
QY 1287 GCTCTCAAGACGCTGCTGCGCTCGAGCTGCGTTCACCCAGCAGCGGTGCTGTGC 1346  
Db 3697 CCGGAGAGGCCCGGCTGCTGCGGCGCCCGCAGGAGGACACAGACCCCGCTCG 3756  
QY 1347 CCGGAGAGGCCCGGCTGCTGCGGCGCCCGCAGGAGGACACAGACCCCGCTCG 1406  
Db 3757 CTTGCTGAGCTGCTCCGCGCAGCAGCAGCCCTGCGAGGTGTACGGCTTCTGTCGGGC 3816  
QY 1407 CTTGCTGAGCTGCTCCGCGCAGCAGCAGCCCTGCGAGGTGTACGGCTTCTGTCGGGC 1466  
Db 3817 CTGCTGCGCGGCTGCTGCGGCGCTCTGCGGCTCCAGGACACAGACCGCGCTT 3876  
QY 1467 CTGCTGCGCGGCTGCTGCGGCGCTCTGCGGCTCCAGGACACAGACCGCGCTT 1526  
Db 3877 CCTCAGGAACACCAAGATTCTCCTCGGGAAGCATGCCAAGCTCTCGCTCAGGA 3936  
QY 1527 CCTCAGGAACACCAAGATTCTCCTCGGGAAGCATGCCAAGCTCTCGCTCAGGA 1586  
Db 3937 GCTGACGTGGAAGATGAGCGTTCGGGACTGCGCTTGGCTGCGCAGGAGCCCGAG 3990  
QY 1587 GCTGACGTGGAAGATGAGCGTTCGGGACTGCGCTTGGCTGCGCAGGAGCCCGAG 1640

## RESULT 13

ID US-09-042-460-1 STANDARD; DNA; UNC; 3496 BP.

AC xxxxxx

DF

Sequence 1, Application US/09042460

Sequence 1, Application US/09042460

GENERAL INFORMATION:

APPLICANT: Morin, Gregg B.

APPLICANT: Allsopp, Richard

APPLICANT: Depinho, Ronald

APPLICANT: Greenberg, Roger

TITLE OF INVENTION: Mouse Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 101

CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,460

FILING DATE: 16-MAR-1998

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

QY

Db

QY

Db

QY

Db

CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,549  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,584  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/979,742  
CC FILING DATE: 26-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-003110US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3496 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY:  
CC LOCATION: 1..3496  
CC OTHER INFORMATION: /note= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT) cDNA clone"  
CC FEATURE:  
CC NAME/KEY: misc\_feature  
CC LOCATION: 10..3435  
CC OTHER INFORMATION: /note= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT) cDNA"  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 39..3404  
CC OTHER INFORMATION: /product= "mouse telomerase reverse  
CC OTHER INFORMATION: transcriptase (mTRT)"  
CC SEQUENCE 3496 BP; 762 A; 998 C; 928 G; 808 T; 0 OTHER.  
SQ  
Query Match 23.6%; Score 453; DB 23; Length 3496;  
Best Local Similarity 64.3%; Pred. No. 0.00e+00;  
Matches 1216; Conservative 0; Mismatches 592; Indels 84; Gaps 3;  
Db 23 CCGGGCTTGGAGCACAATGACCCGCGCTCTCGTTGCGCCGCGGTGCGCTCTCTGCTGCG 82  
QY 51 CCGGGCTTGGAGCACAATGACCCGCGCTCTCGTTGCGCCGCGGTGCGCTCTCTGCTGCG 110  
Db 83 CAGCCGATACGGGAGGTGTGGCGCTGGCAACCTTGTGGCGGCGCTGGGGCCCGGAGG 142

QY 111 CAGCACTACCGGAGGTGCTGCCGCTGGCCACAGTTCTGTGGGGCCCTGGGGCCCCAGGG 170  
Db 143 CAGCGGGCTGTGCAACCCGGGAGCCCGAAGATCTACCGCATTTGGTTGGCCCAATGCTT 202  
QY 171 CTGGCGGCTGTGAGCGCGGGAGCCCGCGGCTTTCCGCGGCTGTGGTGGCCAGTGCT 230  
Db 203 AGTGTGATGCACTGGGGCTCACAGCCTCCACCTGCCAGCTTTCCTTCCACCAAGGTGTC 262  
QY 231 GGTGTGCGTGGCTTGGGACGACGGCCCGCCCGCCCGCCCTCTCTTCCGCCAGGTGTC 290  
Db 263 ATCCCTGAAGAGCTGTGGCAGGTTGTGCAGAGACTCTCGAGCGCAACGAGAGAAA 322  
QY 291 CTGCTGAAGAGCTGTGGCCCGAGTCTGCAGAGGCTGTGCGAGCGGGCGGCAAGAA 350  
Db 323 CGTGTGCTTTTGGCTTTGAGCTGCTTAAAGAGCCAGAGCGGGCTTCCCATGGCCCTT 382  
QY 351 CGTGTGCGCTTTCGCTTTCGGGCTGCTGGAGGGGCGCGGGGCGCCCGCCGAGCGCTT 410  
Db 383 CACTAGTAGCTGCTAGCTACTTGGCCCAACACTGTTATTGAGACCTTGCCTGTCACTGG 442  
QY 411 CACCAACAGCGTGGCAGCTACTTGCCTCAACACAGGTTGACCGACGCTGCGGGGGAGCG 470  
Db 443 TGCATGATGCTACTTGTAGCCGAGTGGGCGGACAGCCTGCTGGTCTACCTGCTGGCACA 502  
QY 471 GCGTGGGGCTGTGCTGCGCGGCTGGCGACGACGCTGCTGCTACCTGCTGGCAG 530  
Db 503 CTGTGCTTTTATCTTTGCTGGCCCCCAGCTGTGCTTACCAAGTGTGGTGTCCCT 562  
QY 531 CTGCGCGCTTTTGTGTGGTGGCTTCCAGCTGCGCTTACCAGGTGTGCGGGCGCGCCT 590  
Db 563 GTACCAATTTTGCACACGATATCTGGCCCTCTGTGCTCCGCTAGTTACAGGCCAC 622  
QY 591 GTACAGCTGCGGCTGCACTAGCGCGCGCGCCCGCCACACGCTAGTGGACCCCGAAG 650  
Db 623 CCGACCCCTGGCGAGGAATTTCACTTAACCTTAGGTTCTTACAACAGATCAAGAGCAGTAG 682  
QY 651 CGCTGGGATGCG-----AACGGGCTTGAACCATAG 683  
Db 683 TCGCCAGGAAGACCGAACCCTTGGCTTGCCTTCCATCTCGAGGTACAAAGAGCATCTGAG 742  
QY 684 CGTCAGGAGGCGGGGTCCTTGGGCTGCGAGCCCGGCTGGAGGAGCGCGGGGG 743  
Db 743 TCTACCACTACAGTGTGCTTCACTTAAGAGCCAGATGCTATCTCTTAAAGAAAGGT 802  
QY 744 CAGTGCCACGCAAGTCTGCGCTTGGCCAAAGAGCGCCAGGCTGCGCTGCCCTGAGCC 803  
Db 803 GGAGGAGGACCCACAGGAGGTGCTTACCAACCCCATCAGSCAAATCATGGTGGCCAAAG 862  
QY 804 GGAGGAGGACCCGTTGGGAGGGTCTTGGGCGCCAGCGGCGGAGCGGTGGACCCAG 863  
Db 863 TCCTGCTCGTCCCGGAGGTGCTTACTTCAAGAGAAAGATTGCTTCTTAAAGAAAGGT 922  
QY 864 TGACCGTGTGTTCTGTGGTGTACCTTGCAGACGCCCGCGAAGACCACTCTTTGGA 923  
Db 923 GTCGTACCTGAGTCTCTTGG-----TCGGTGTGCTGTAACACAGCCCGAG 970  
QY 924 GGGTGCCTCTGTGACGCGGCCACTTCCACCCATCCGTGGGCGCCCGCCAGCACCGCGG 983  
Db 971 TCCACATCTGTGCTCACCAACCGCCGCAAAATGCTTTTCAGCTCAGGCCCATTTATGA 1030  
QY 984 CCCCCATCCAGATCGCGGCCACACAGTCTCTGGGACAGCGCTTGTCCCGGGGTAGCG 1043  
Db 1031 GACCAAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGGGTCTTAAACCCCTCATTTCT 1090  
QY 1044 CGAGACCAAGCATTTCTTCTACTCTCAGCGGACAGAGGAGCTGCGCGCCCTCTTCT 1103  
Db 1091 ACTCAGCAACCTTCAGCCTACTTGTGCTGGGGCCAGAGAGCTGTGGAGATCATCTTTCT 1150  
QY 1104 ACTAGCTCTCTGAGGCGCCGCTGACTTGGGCTCGGAGGCTGTGGAGACCATCTTTCT 1163  
Db 1151 GGGCTCAGGCGCTAGGACATCAGGACCATCTGCAAGGACACACCTCTATCGCGTCGATA 1210  
QY 1164 GGTTCACGGCCCTGGATGCGGAGGACTTCCCGCGAGGTGTCGCCCGCCTGCCCCAGCGCTA 1223

Db 1211 CTGCGAGATGGGCCCCCTGTTTCCAAACAGCTGCTGGTGAACCATCGAGAGTGCCCAATATGT- 1270  
QY 1224 CTGCAAAATGGGCCCCCTGTTTCTGGAGCTGCTTGGAAACACCGCGAGTGCCCTACGG 1283  
Db 1271 CAGACTCTCTCAGGTCACATTGCGAGTTTTCGAACAGCAAAACCAAGGTGACAGATGCTTT 1330  
QY 1284 GGTGCTCTCAAGAGCACTGCCCGCTGCGAGTGCCTTACCACAGAGCGGTGCTGTG 1343  
Db 1331-----GAACACACGACCCAC 1345  
QY 1344 TGCCCGGAGAGCCCGAGGGCTCTGTGGCGCCCCCGAGGAGGAGGACACAGACCCCG 1403  
Db 1346 GCACCTCATGATTTGCTTCGCGCTGCACAGAGTCCCTTGGCAGGTATATGTTTCTTTCG 1405  
QY 1404 TCGCTGTGTGAGCTGCTCCGCCACGACAGCAGCCCTTGGCAGGTGTACGGCTTCGTGCG 1463  
Db 1406 GGCTGTCTCTGCAAGGTGTGTCTGTCTGTGGGTACCAAGGACACATGAGCGCG 1465  
QY 1464 GGCTGTCTGCGCGGCTGTGTGCCCGAGGCTCTGGGGCTTCCAGGCACAAACGACCGCG 1523  
Db 1466 CTTCTTTAAGAACTTAAAGAGTTTCATCTGTTGGGAAATACGCAAGCTATCACTGCA 1525  
QY 1524 CTTCTCAGGAACACCAAGAGTTTCATCTCCTTGGGAGCATGCCAAGCTCTCGCTGCA 1583  
Db 1526 GGAATGATGTGGAAGATGAAAGTAGAGATTGCCACTGGCTCCGACAGCCCGGGGAA 1585  
QY 1584 GGAGCTACGCTGGAAGATGAGCGTGGGAGTGGCTTGGCTGGCAGAGCCCGAGGGT 1643  
Db 1586 GGACCGTGTCCCCCTGCAGAGCACCGTCTGAGGAGAGGATCTGCTAGGTTCCTGTT 1645  
QY 1644 TGGCTGTGTTCCGCGCCAGAGCACCTCTGCGTGGAGGATCTGCGCAAGTTCTCTGCA 1703  
Db 1646 CTGCTGATGACACATACGTGGTACAGCTGCTTAGGTTCATTTTACATCAGAGAG 1705  
QY 1704 CTGCTGATGAGTGTACGCTGCTGAGCTGCTCAGGTCTTCTTTATGTCAGGAGAC 1763  
Db 1706 CACATTCAGGAAGAACAGGCTTCTTCTTCCGTAAGAGTGTGTGGAGCAAGTGCAGAG 1765  
QY 1764 CACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAG 1823  
Db 1766 CATTGGAGTCAAGCAACACTTGAGAGATGCGGCTACGGAGCTGTACAGAGGAGGT 1825  
QY 1824 CATTGGAATCAGACAGCACTTGAAGAGGTGCACTGCGGAGGCTGTGGAAGAGAGGT 1883  
Db 1826 CAGCATCACCAGCACACCTGGCTAGCCATGC 1857  
QY 1884 CAGCAGATCGGAAGCCAGGCCCGCCCTGC 1915

RESULT 14  
ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.  
AC xxxxx  
DT

DE Sequence 36, Application US/09026981  
CC Sequence 36, Application US/09026981  
CC GENERAL INFORMATION:  
CC APPLICANT: Counter, Christopher M.  
CC APPLICANT: Meyerson, Matthew  
CC APPLICANT: Weinberg, Robert A.  
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
CC City: Lexington  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02173  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/026,981  
CC FILING DATE: 20-FEB-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/064,322  
CC FILING DATE: 30-OCT-1997  
CC APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/055,762  
CC FILING DATE: 14-AUG-1997  
CC APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/054,549  
CC FILING DATE: 01-AUG-1997  
CC APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/047,151  
CC FILING DATE: 20-MAY-1997  
CC APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/038,750  
CC FILING DATE: 20-FEB-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Granahan, Patricia  
CC REGISTRATION NUMBER: 32,227  
CC REFERENCE/DOCKET NUMBER: WH197-11p4AM  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 781-861-6240  
CC INFORMATION FOR SEQ ID NO: 36:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3346 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
SQ SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 14.6%; Score 281; DB 23; Length 3346;  
Best Local Similarity 100.0%; Pred. No. 2,85e-234;  
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 971 GGGTTGGCTGTCTCGGCGCGAGACACCGTCTCGTGAGGAGATCTTGGCCAAAGTTCC 1030  
Qy 1640 GGGTTGGCTGTCTCGGCGCGAGACACCGTCTCGTGAGGAGATCTTGGCCAAAGTTCC 1699  
Db 1031 TGCACCTGGCTGATGATGTGTACGTCGTCGAGCTGTCTTCTTTTATGTACAGG 1090  
Qy 1700 TGCACCTGGCTGATGATGTGTACGTCGTCGAGCTGTCTTCTTTTATGTACAGG 1759  
Db 1091 AGACACCGTTTCAAAAGACAGCGCTCTTTTCTACGGAGAGTGTCTGGACAAAGTTGC 1150  
Qy 1760 AGACACCGTTTCAAAAGACAGCGCTCTTTTCTACGGAGAGTGTCTGGACAAAGTTGC 1819  
Db 1151 AAAGCATTGGAATCAGACAGCACTTGAAGAGGTGAGCTGGGGAGCTGTCTGGGAAGCAG 1210  
Qy 1820 AAAGCATTGGAATCAGACAGCACTTGAAGAGGTGAGCTGGGGAGCTGTCTGGGAAGCAG 1879  
Db 1211 AGGTGAGGAGCAGTGGGAAGCCAGCGCCCGCTGTGACG 1251  
Qy 1880 AGGTGAGGAGCAGTGGGAAGCCAGCGCCCGCTGTGACG 1920

RESULT 15  
ID US-08-911-312-52 STANDARD; DNA; UNC; 535 BP.  
AC xxxxxx

DE Sequence 52, Application US/08911312  
CC Sequence 52, Application US/08911312  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William  
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 170  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/911,312  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Einhorn, Gregory P.  
CC REGISTRATION NUMBER: 38,440  
CC REFERENCE/DOCKET NUMBER: 015389-0025000US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 52:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 535 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
SQ SEQUENCE 535 BP; 105 A; 152 C; 173 G; 105 T; 0 OTHER.

Query Match 13.5%; Score 259; DB 22; Length 535;  
Best Local Similarity 100.0%; Pred. No. 6.37e-213;  
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 AGAGACCGTCTCGGTGAGGAGATCTTGGCCAAAGTTCTTCTGACCTGGCTGATGATGTGTA 60  
Qy 1662 AGAGACCGTCTCGGTGAGGAGATCTTGGCCAAAGTTCTTCTGACCTGGCTGATGATGTGTA 1721  
Db 61 CGTGTGAGCTGTCTGAGTCTTTCTTTATGTCTACGAGACCACTTCTTCAAAAGACAG 120  
Qy 1722 GGTGTGAGCTGTCTGAGTCTTTCTTTATGTCTACGAGACCACTTCTTCAAAAGACAG 1781  
Db 121 GCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAAAGCATTTGGAATCAGACAGCA 180  
Qy 1782 GCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAAAGCATTTGGAATCAGACAGCA 1841  
Db 181 CTTGAAGAGGTGAGCTGCGGAGAGTGTCTCGGAAGAGAGGTCTCAGGACGATCGGGAAGC 240

Qy 1842 CTTGAAGAGGGTGCAGCTGCGGGAGCTGTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGC 1901

Db 241 CAGGCCCGCCCTGCTGACG 259

Qy 1902 CAGGCCCGCCCTGCTGACG 1920

Search completed: Thu Dec 24 16:09:24 1998  
Job time : 3547 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*

W P S R E F L I (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 15:07:06 1998; MasPar time 156.09 Seconds  
Tabular output not generated. 573.790 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1-1920) from US08951733.seq (1 of 2)  
Perfect Score: 1920  
N.A. Sequence: 1 CACCGGTCGGGCGAGCGCTG.....CCAGGCGCGCTGCTGACG 1920  
Comp: GTGCGCAGGCGCGTCGCGAC.....GGTCCGGCGGCGAGCTGC

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 8.977; Variance 5.692; scale 1.577

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	67	3.5	7218	1	US-08-232- Sequence 14, Applicati	1.19e-24
2	44	2.3	7218	1	US-08-232- Sequence 14, Applicati	2.87e-11
3	34	1.8	215	1	US-08-238- Sequence 5, Applicatio	7.29e-06
4	35	1.8	215	1	US-08-238- Sequence 5, Applicatio	2.19e-06
5	26	1.4	74	2	PCT-US95-1 Sequence 100, Applicat	7.10e-02
6	26	1.4	74	2	PCT-US95-1 Sequence 94, Applicati	7.10e-02
7	26	1.4	81	2	PCT-US95-1 Sequence 92, Applicati	7.10e-02
8	26	1.4	81	2	PCT-US95-1 Sequence 98, Applicati	7.10e-02
9	26	1.4	82	2	PCT-US95-1 Sequence 97, Applicati	7.10e-02
10	26	1.4	82	2	PCT-US95-1 Sequence 97, Applicati	7.10e-02
11	25	1.3	66	1	US-08-471- Sequence 144, Applicat	2.09e-01
12	25	1.3	68	1	US-07-977- Sequence 243, Applicat	2.09e-01
13	25	1.3	69	1	US-08-471- Sequence 142, Applicat	2.09e-01
14	25	1.3	74	2	PCT-US95-1 Sequence 94, Applicati	2.09e-01
15	25	1.3	74	2	PCT-US95-1 Sequence 100, Applicat	2.09e-01
16	25	1.3	75	2	PCT-US95-1 Sequence 99, Applicati	2.09e-01
17	25	1.3	75	2	PCT-US95-1 Sequence 99, Applicati	2.09e-01
18	25	1.3	81	2	PCT-US95-1 Sequence 98, Applicati	2.09e-01
19	25	1.3	81	2	PCT-US95-1 Sequence 98, Applicati	2.09e-01

c	20	25	1.3	242	1	US-08-273- Sequence 1, Applicatio	2.09e-01
c	21	25	1.3	1004	2	PCT-US95-0 Sequence 7, Applicatio	2.09e-01
c	22	25	1.3	1386	2	PCT-US95-0 Sequence 2, Applicatio	2.09e-01
c	23	25	1.3	1611	2	PCT-US93-0 Sequence 3, Applicatio	2.09e-01
c	24	23	1.2	59	2	PCT-US95-1 Sequence 95, Applicati	1.70e+00
c	25	24	1.2	65	1	US-08-471- Sequence 145, Applicati	6.02e-01
c	26	24	1.2	66	2	PCT-US95-1 Sequence 93, Applicati	6.02e-01
c	27	23	1.2	66	1	US-08-471- Sequence 144, Applicat	1.70e+00
c	28	24	1.2	68	1	US-08-471- Sequence 143, Applicat	6.02e-01
c	29	23	1.2	69	1	US-08-471- Sequence 142, Applicat	1.70e+00
c	30	23	1.2	84	1	US-08-133- Sequence 120, Applicati	1.70e+00
c	31	23	1.2	84	1	US-08-209- Sequence 23, Applicati	1.70e+00
c	32	23	1.2	84	2	PCT-US94-0 Sequence 25, Applicati	1.70e+00
c	33	23	1.2	84	1	US-08-300- Sequence 25, Applicati	1.70e+00
c	34	23	1.2	225	1	US-07-807- Sequence 15, Applicati	1.70e+00
c	35	23	1.2	225	1	US-08-299- Sequence 18, Applicati	1.70e+00
c	36	23	1.2	1404	3	5480796-8 Patent No. 5480796	1.70e+00
c	37	23	1.2	1404	3	5171840-8 Patent No. 5171840	1.70e+00
c	38	23	1.2	1640	1	US-07-807- Sequence 11, Applicati	1.70e+00
c	39	23	1.2	1640	1	US-08-299- Sequence 11, Applicati	1.70e+00
c	40	23	1.2	2061	3	5480796-1 Patent No. 5480796	1.70e+00
c	41	23	1.2	2061	3	5171840-1 Patent No. 5171840	1.70e+00
c	42	23	1.2	2253	3	5457037-2 Patent No. 5457037	6.02e-01
c	43	24	1.2	2454	1	US-07-872- Sequence 1, Applicatio	1.70e+00
c	44	23	1.2	2625	3	5457037-4 Patent No. 5457037	1.70e+00
c	45	23	1.2	3336	3	5457037-1 Patent No. 5457037	1.70e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx  
DT  
DE Sequence 14, Application US/08232463  
CC Sequence 14, Application US/08232463  
CC Patent No. 5670367  
CC GENERAL INFORMATION:  
CC APPLICANT: DORNER, F.  
CC APPLICANT: SCHEIFLINGER, F.  
CC APPLICANT: FALKNER, F. G.  
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 1800 Diagonal Road, Suite 500  
CC CITY: Alexandria  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22313-0299  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/232.463  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/935.313  
CC FILING DATE:  
CC APPLICATION NUMBER: EP 91 114 300.6  
CC FILING DATE: 26-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703)836-9300  
CC TELEFAX: (703)683-4109  
CC TELEX: 899149  
CC INFORMATION FOR SEQ ID NO: 14:









AC xxxxxx  
DE Sequence 98, Application PC/TUS9511934  
CC Sequence 98, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 98:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 81 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.  
Query Match 1.4%; Score 26; DB 2; Length 81;  
Best Local Similarity 14.3%; Pred. No. 7.10e-02;  
Matches 10; Conservative 19; Mismatches 41; Indels 0; Gaps 0;  
Db 1 GTTGTCTAGANNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 60  
Cp 1515 GTTGTCTAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAG 1456  
Db 61 VNNVNNVNNV 70  
Cp 1455 GCGGTACACC 1446  
RESULT 9  
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.  
AC xxxxxx  
DE Sequence 97, Application PC/TUS9511934  
CC Sequence 97, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 97:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 82 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 82 BP; 1 A; 2 C; 10 G; 8 T; 61 OTHER.  
Query Match 1.4%; Score 26; DB 2; Length 82;  
Best Local Similarity 15.2%; Pred. No. 7.10e-02;  
Matches 12; Conservative 20; Mismatches 47; Indels 0; Gaps 0;  
Db 2 TGTCTCTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 61  
Cp 158 TGGGGCCCCAGGCTGGCTGGTGCAGCGCGGACCGCGGCTTCCGCGGCTGG 217  
Db 62 NNN 80  
Cp 218 TGGCCCACTGCTGGTGTG 236  
RESULT 10  
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.  
AC xxxxxx  
DE Sequence 97, Application PC/TUS9511934  
CC Sequence 97, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 97:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 82 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 82 BP; 1 A; 2 C; 10 G; 8 T; 61 OTHER.  
Query Match 1.4%; Score 26; DB 2; Length 82;  
Best Local Similarity 15.2%; Pred. No. 7.10e-02;  
Matches 12; Conservative 20; Mismatches 47; Indels 0; Gaps 0;  
Db 2 TGTCTCTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 61  
Cp 158 TGGGGCCCCAGGCTGGCTGGTGCAGCGCGGACCGCGGCTTCCGCGGCTGG 217  
Db 62 NNN 80  
Cp 218 TGGCCCACTGCTGGTGTG 236  
RESULT 10  
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.  
AC xxxxxx  
DE Sequence 97, Application PC/TUS9511934  
CC Sequence 97, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864



CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864

```
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 100:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 74 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.

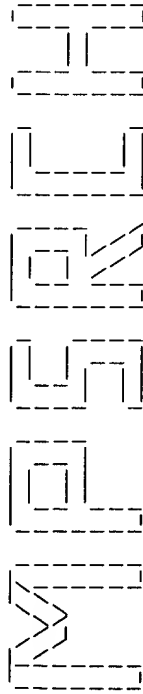
Query Match      1.3%; Score 25; DB 2; Length 74;
Best Local Similarity 7.7%; Pred. No. 2.09e-01;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db      6 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 65
Qy      945 CCACTCCCAACCATCCGTGGCGGCCAGCACACGCGGGCCCCCATCCACATCGCGGCC 1004

Db      66 ACCAC 70
Qy      1005 ACCAC 1009
```

Search completed: Thu Dec 24 15:09:57 1998  
Job time : 171 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties  
Run on: Thu Dec 24 16:09:43 1998; MasPar time 2754.39 Seconds  
720.892 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-951-733-13  
Description: (1920-2820) from US08951733.seq (2 of 2)  
Perfect Score: 901  
N.A. Sequence: 1920 GTCCAGACTCCGCTTCATCC.....AGCCCTGGTGGCAGCGGT 2820  
Comp: CAGGCTGAGCGCAAGTAGG.....TCCGGGACCCACCGTGCCGA

Scoring table: TABLE default  
Gap open 30; Gap extend 1  
Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_vi  
Database: genbank107  
15:gb\_bal 16:gb\_ba2 17:gb\_htg 18:gb\_in 19:gb\_om 20:gb\_ov  
21:gb\_pat 22:gb\_ph 23:gb\_pl1 24:gb\_pl2 25:gb\_pri  
26:gb\_pr2 27:gb\_pr3 28:gb\_ro 29:gb\_st 30:gb\_sts 31:gb\_sy  
32:gb\_un 33:gb\_vi

Statistics: Mean 10.599; Variance 5.047; scale 2.100

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	DB ID	Description	Pred. No.
1	901	100.0	4015	27	AF015950 Homo sapiens telomeras	0.00e+00
2	901	100.0	4027	27	AF018167 Homo sapiens telomeras	0.00e+00
3	901	100.0	8960	31	AF043739 Synthetic construct hu	0.00e+00
4	369	41.0	3369	28	AF073311 Mus musculus telomeras	7.56e-290
5	369	41.0	3426	28	AF051911 Mus musculus telomeras	7.56e-290
6	39	4.3	7218	21	I66494 Sequence 14 from paten	1.71e-09
7	34	3.8	74371	26	AC005369 Homo sapiens chromosom	3.54e-06
8	30	3.3	215	21	I28278 Sequence 5 from patent	1.16e-03
9	30	3.3	74371	26	AC005369 Homo sapiens chromosom	1.16e-03
10	29	3.2	15807	15	MEU72662 Methylobacterium extor	4.69e-03
11	28	3.1	215	21	I28278 Sequence 5 from patent	1.85e-02
12	26	2.9	5591	15	AF027499 Azotobacter vinelandii	2.66e-01

C	13	25	2.8	1908	28	BRU02553	Rattus rattus protein	9.67e-01
C	14	25	2.8	1912	28	SR1478	3GH13/CL100 pPase-ox	9.67e-01
C	15	25	2.8	1948	28	RNRNADSP	R.norvegicus mRNA for	9.67e-01
C	16	25	2.8	7218	21	I66494	Sequence 14 from paten	9.67e-01
C	17	24	2.7	1690	26	HSRNUA14	H.sapiens mRNA for Uli	3.40e+00
C	18	24	2.7	1789	28	RNHLSHR	R.norvegicus H1SHR MRN	3.40e+00
C	19	24	2.7	2140	15	S68984	eryM-malonyl-coenzyme	3.40e+00
C	20	24	2.7	2140	15	SERVALCOAD	Saccharopolyspora eryt	3.40e+00
C	21	24	2.7	2699	26	AB006713	Homo sapiens mRNA for	3.40e+00
C	22	24	2.7	6462	23	AF044920	Volvox carteri acetola	3.40e+00
C	23	24	2.7	7011	28	RATRBB1	Rat calcium channel al	3.40e+00
C	24	24	2.7	6955	28	AF055477	Rattus norvegicus pore	3.40e+00
C	25	24	2.7	26816	25	HUMLUCAL16	Human cosmid LUCA16, c	3.40e+00
C	26	24	2.7	76199	15	AF040570	Amycolatopsis mediterr	3.40e+00
C	27	24	2.7	92879	17	AC004450	*** SEQUENCING IN PROG	3.40e+00
C	28	24	2.7	125532	17	AC004975	*** SEQUENCING IN PROG	3.40e+00
C	29	24	2.7	189370	15	AF010496	Rhodobacter capsulatus	3.40e+00
C	30	24	2.7	216021	26	HUAC004787	Homo sapiens Chromosom	3.40e+00
C	31	24	2.7	317311	17	AC004085	*** SEQUENCING IN PROG	3.40e+00
C	32	23	2.6	6621	141364	I41364	Sequence 144 from pate	1.16e+01
C	33	23	2.6	6921	I41362	I41362	Sequence 142 from pate	1.16e+01
C	34	23	2.6	8421	I65426	I65426	Sequence 25 from paten	1.16e+01
C	35	23	2.6	8421	I61313	I61313	Sequence 120 from pate	1.16e+01
C	36	23	2.6	1160	15	CGAJ4934	Corynebacterium glutam	1.16e+01
C	37	23	2.6	2281	28	MUSHAL	Mouse histidase locus	1.16e+01
C	38	23	2.6	4984	15	AF051692	Pseudomonas aeruginosa	1.16e+01
C	39	23	2.6	5636	18	AB002270	Bombyx mori L1Bm DNA,	1.16e+01
C	40	23	2.6	1176	33	HPEGENSA	Hepatitis E virus ORF1	1.16e+01
C	41	23	2.6	7200	33	HPCGENOM	Hepatitais E virus com	1.16e+01
C	42	23	2.6	38582	18	CELK12HA	C. elegans cosmid K12H	1.16e+01
C	43	23	2.6	53784	15	AMW223012	Amycolatopsis mediterr	1.16e+01
C	44	23	2.6	76199	15	AF040570	Amycolatopsis mediterr	1.16e+01
C	45	23	2.6	216021	26	HUAC004787	Homo sapiens Chromosom	1.16e+01

ALIGNMENTS

RESULT	1	AF015950	4015 bp	mRNA	PRI	16-AUG-1997
LOCUS						
DEFINITION						
ACCESSION		AF015950				
NID		92330016				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		1 (bases 1 to 4015)				
AUTHORS		Nakamura,T.M., Morin,G.B., Chapman,K.B., Weinrich,S.L., Andrews,W.H., Lingner,J., Harley,C.B. and Cech,T.R.				
TITLE		Telomerase catalytic subunit homologs from fission yeast and human				
JOURNAL		Science 277 (5328), 955-959 (1997)				
MEDLINE		97400623				
REFERENCE		2 (bases 1 to 4015)				
AUTHORS		Morin,G.B.				
TITLE		Direct Submission				
JOURNAL		Submitted (24-JUL-1997) Geron Corporation, 230 Constitution Drive,				
FEATURES		Menlo Park, CA 94025, USA				
source		Location/Qualifiers				
		1. .4015				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/tissue_type="kidney"				
		/map="5p"				
		/dev_stage="embryo"				
		/chromosome="5"				
		1. .4015				
		/gene="hTRT"				
		56.3454				
		/gene="hTRT"				
		/codon_start=1				

```
/product="telomerase reverse transcriptase"
/db_xref="pid:g2330017"
/translation="MPRAPRCRAVRSLRSHRYREVLPLATFVRRLGPGQWRVLVQRGDP
AAFRALVACQVCPVDPMDRPPAPSPFQVSCLELVARVLQRLCERAGKNVLAFGFA
LLDGARGGPEAFVTSVSYLPTNTVDALRGSGAWGLLRVGDVLLVHLALCALFV
LVAPCAQVCGPPLYOLGAATQAPRPHASGPRRLGRCERANHSVREAGVPLGPA
PGARRGGSASRLPLPKPRRGAAPEPTPVQGSWAHPGRTGRPSDRGFCVWSPA
RPAEATSLGALSCTRHSHPSVGRQHAGPPSTSRPPRPWDTPCPVYAEATHFELYS
SGDKEQURPSFLSSRLSUTGARRLVETIFLGSRPMPGTPRRLPLPQRYWMRPL
FLELLGHNACQPYGLVLTHTCPLRAAVTPAAGCAREKPGQSVAAPEEDTPRRVLQ
LLRHSGSPWQVYGFVRACLRRLVPPGLWGRSHNERRLNKKFISLGRKAKLSLOEL
TKMSVRDCAWLRSPGCVGCPVPAEHLREILLAKFLHMLSVVVELLSFKSVYET
TFQKRLFFYKRSVWSKLSIGIRQHLKRYQLRELSAEVROHREARPALTSRLRFI
PKDGLRPIVNDIVGARTFRREKRAELTSRVKALFVSLNTERAKRPGLLGASVLG
LDDIHRARTFVLRVADPPPELYFYKVDVTGAYDTPQDRLTEVIATSIKQNTYC
VRYAYVQKAHGHVRFKAFKSHVSTLDQPMKQFVAHLQETSPLRDADVIVQSSSL
NEASSGLFVDFMCHHVRKFSVQVQOGIPQGSILSTLSCLYGDMENKLEAG
IRRDGLLRVDDFLVTPHLLTHAKTFLRLTVRGVEYGVNLRKTVNFPVEDEAL
GCTAFVOMPAHGLFPWCGLLDTLRTLEVQSDYSYARTSIRASLTFRNGFKAGRNRR
KLFGLVRLKCHSLFDPWGLVNSLTQVCTNIYKILLQAIREFHACVQLPFPFQVQKNPT
FYFLVSDTASLQYSLKNAKNGMSLGAAGPLSEAVQWLCHQAFLLKLRHRTV
FVPLGLSLRTAQQLSRKLPGLTITALEAANPALPSDFKLTILD"
BASE COUNT 663 a 1363 c 1275 g 714 t
ORIGIN

Query Match 100.0%; Score 901; DB 27; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1909 GTCCAGACTCGCTTCATCCCAAGCTGACGGCTGCGGCGCATGTGACATGGACTA 1968
Qy 1920 GTCCAGACTCGCTTCATCCCAAGCTGACGGCTGCGGCGCATGTGACATGGACTA 1979

Db 1969 CGTGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGGCGCTCTCACTCGAGGGT 2028
Qy 1980 CGTGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCGGCGCTCTCACTCGAGGGT 2039

Db 2029 GAAGCACTGTTACGCTGCTCACTACGAGCGGCGGCGGCGGCGCTCTCGGCGGC 2088
Qy 2040 GAAGCACTGTTACGCTGCTCACTACGAGCGGCGGCGGCGGCGCTCTCGGCGGC 2099

Db 2089 CTCTGTGCTGGCGCTGAGCATATCCAGAGGCTGCGGCGACCTTCGTGCTGCGTGGCG 2148
Qy 2100 CTCTGTGCTGGCGCTGAGCATATCCAGAGGCTGCGGCGACCTTCGTGCTGCGTGGCG 2159

Db 2149 GGCCAGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGGATGACGGCGCGGTACGA 2208
Qy 2160 GGCCAGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGGATGACGGCGCGGTACGA 2219

Db 2209 CACCATCCCCAGACAGGCTCAGGAGGTGATCGCCAGCATCATCAACCCCAAGACAC 2268
Qy 2220 CACCATCCCCAGACAGGCTCAGGAGGTGATCGCCAGCATCATCAACCCCAAGACAC 2279

Db 2269 GTACTGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
Qy 2280 GTACTGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339

Db 2329 CTTCAAGAGCCACGCTCTTACCTTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGTCG 2388
Qy 2340 CTTCAAGAGCCACGCTCTTACCTTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGTCG 2399

Db 2389 TCACCTGCAGAGACCCGCTGAGGATGCGGTGCTCATGACAGAGAGCTCCTCCCT 2448
Qy 2400 TCACCTGCAGAGACCCGCTGAGGATGCGGTGCTCATGACAGAGAGCTCCTCCCT 2459

Db 2449 GAATGAGCCAGCAGTGCCTCTTCAGAGCTTCTTACGCTTCATGTCGACACGCGGT 2508
Qy 2460 GAATGAGCCAGCAGTGCCTCTTCAGAGCTTCTTACGCTTCATGTCGACACGCGGT 2519

Db 2509 GCGCATCAGGGGCAAGTCTTACGCTTCAGTGCAGGGGATCCCGCAGGCGTCCATCTCTC 2568
Qy 2520 GCGCATCAGGGGCAAGTCTTACGCTTCAGTGCAGGGGATCCCGCAGGCGTCCATCTCTC 2579
```

```
Db 2569 CACGCTCTCTGAGCCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 2628
Qy 2580 CACGCTCTCTGAGCCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 2639

Db 2629 GCGGAGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2688
Qy 2640 GCGGAGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2699

Db 2689 CCACGCGAAAAACCTTCTCAGGACCCCTGCTCCGAGGTGCTCCCTGAGTATGGCTGCGTGGT 2748
Qy 2700 CCACGCGAAAAACCTTCTCAGGACCCCTGCTCCGAGGTGCTCCCTGAGTATGGCTGCGTGGT 2759

Db 2749 GAACTTCGGAAGACAGTGTGTAACCTTCCCTGTAGAGACGAGGCGCTGCTGCTGCTGCTGCTG 2808
Qy 2760 GAACTTCGGAAGACAGTGTGTAACCTTCCCTGTAGAGACGAGGCGCTGCTGCTGCTGCTGCTG 2819

Db 2809 T 2809
Qy 2820 T 2820

RESULT 2
LOCUS AF018167 4027 bp mRNA PRI 29-AUG-1997
DEFINITION Homo sapiens telomerase catalytic subunit (hEST2) mRNA, complete cds.
ACCESSION AF018167
NID 92347128
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4027)
AUTHORS Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q., Bacchetti, S., Haber, D.A. and Weinberg, R.A.
TITLE hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization
JOURNAL Cell 90 (4), 785-795 (1997)
MEDLINE 97433088
REFERENCE 2 (bases 1 to 4027)
AUTHORS Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J., Liu, Q., Bacchetti, S., Haber, D.A. and Weinberg, R.A.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142, USA
FEATURES
source
1..4027
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p15.33"
1..4027
/note="human Ever Shorter Telomeres 2"
57..3455
/gene="hEST2"
/gene="hEST2"
/note="member of reverse transcriptase class"
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g2347129"
/translation="MPRAPRCRAVRSLRSHRYREVLPLATFVRRLGPGQWRVLVQRGDP
AAFRALVACQVCPVDPMDRPPAPSPFQVSCLELVARVLQRLCERAGKNVLAFGFA
LLDGARGGPEAFVTSVSYLPTNTVDALRGSGAWGLLRVGDVLLVHLALCALFV
LVAPCAQVCGPPLYOLGAATQAPRPHASGPRRLGRCERANHSVREAGVPLGPA
PGARRGGSASRLPLPKPRRGAAPEPTPVQGSWAHPGRTGRPSDRGFCVWSPA
RPAEATSLGALSCTRHSHPSVGRQHAGPPSTSRPPRPWDTPCPVYAEATHFELYS
SGDKEQURPSFLSSRLSUTGARRLVETIFLGSRPMPGTPRRLPLPQRYWMRPL
FLELLGHNACQPYGLVLTHTCPLRAAVTPAAGCAREKPGQSVAAPEEDTPRRVLQ
LLRHSGSPWQVYGFVRACLRRLVPPGLWGRSHNERRLNKKFISLGRKAKLSLOEL
TKMSVRDCAWLRSPGCVGCPVPAEHLREILLAKFLHMLSVVVELLSFKSVYET
TFQKRLFFYKRSVWSKLSIGIRQHLKRYQLRELSAEVROHREARPALTSRLRFI
PKDGLRPIVNDIVGARTFRREKRAELTSRVKALFVSLNTERAKRPGLLGASVLG
LDDIHRARTFVLRVADPPPELYFYKVDVTGAYDTPQDRLTEVIATSIKQNTYC
VRYAYVQKAHGHVRFKAFKSHVSTLDQPMKQFVAHLQETSPLRDADVIVQSSSL
NEASSGLFVDFMCHHVRKFSVQVQOGIPQGSILSTLSCLYGDMENKLEAG
IRRDGLLRVDDFLVTPHLLTHAKTFLRLTVRGVEYGVNLRKTVNFPVEDEAL
GCTAFVOMPAHGLFPWCGLLDTLRTLEVQSDYSYARTSIRASLTFRNGFKAGRNRR
KLFGLVRLKCHSLFDPWGLVNSLTQVCTNIYKILLQAIREFHACVQLPFPFQVQKNPT
FYFLVSDTASLQYSLKNAKNGMSLGAAGPLSEAVQWLCHQAFLLKLRHRTV
FVPLGLSLRTAQQLSRKLPGLTITALEAANPALPSDFKLTILD"
LURQSSPMQVYGFVRACLRRLVPPGLWGRSHNERRLNKKFISLGRKAKLSLOEL
```



TKMSVRCGAWLRSPGCVPAAEHRLREEILAKFLHMLSVYVVELLSRFFVTET  
TFQKRLFFYKRSWKLQSIGIRQHLKRVOLRELSAEVROHREARPAALLTSRLRFI  
PKDGLPIVNDYVVGARFREREKRAERLTSVKALFVNLTERARPGLLGASVIG  
LDDIHRAMRTVLRRAQDPPPELIFVKVDVTGAYDIPQDRLTEVIASTIKPQNTYC  
VRYAVVQAAHGHVRAKFSHSTLTDLPYMRQFVAHQETSPLEDAVVIIPQSSSL  
NEASSGLFVDFRMCCHAVIRKGSVQCGIPQGSILSLCSCYGDGMENKLFAG  
IRDDGLLLRVDLFLVTHLTHAKTFLRTLVRGVEYGVNLRKTVVNFVEDEAL  
GGTAFVQMPAHGFLPWCGLLDLDTLEVDOSDSYXARTSIRASLTNRGCFKAGNRWR  
KLFGLVRLKCHSLFLDLQVNSLQVCTNIYKILLQNYRHFACVQLQPHQQVWKAPT  
FFURVLSDTASLFLSILKAKNAGMSIGKAAGALPSEAVQWICHQAFLLKLRRHRT  
YVPLGLSRLTAQOLSKRLPGTLTLTALEAANPALPSDFKTLID"

BASE COUNT 674 a 1361 c 1277 g 715 t

Query Match 100.0%; Score 901; DB 27; Length 4027;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1910 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTGCGCGCATTTGTGAACATGGACTA 1969  
QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGCTGCGCGCATTTGTGAACATGGACTA 1979  
Db 1970 CGTCGTGGGACCCAGAACGTTCCGACAGAAAAGGCGCGAGGCTCTCACCTCGAGGGT 2029  
QY 1980 CGTCGTGGGACCCAGAACGTTCCGACAGAAAAGGCGCGAGGCTCTCACCTCGAGGGT 2039  
Db 2030 GAAGGCACTGTTAGCGTGCTCAACTACGAGCGGGCGGCGCCCGGCTCTCGGGCGC 2089  
QY 2040 GAAGGCACTGTTAGCGTGCTCAACTACGAGCGGGCGGCGCCCGGCTCTCGGGCGC 2099  
Db 2090 CTCTGTGCTGGCGCTGACGATATCCACAGGCGCTGGCGACCTTCGTGCTGCTGTCG 2149  
QY 2100 CTCTGTGCTGGCGCTGACGATATCCACAGGCGCTGGCGACCTTCGTGCTGCTGTCG 2159  
Db 2150 GGCCACGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGTGACGGCGCGTACGA 2209  
QY 2160 GGCCACGAGCCCGCGCTGAGCTGTACTTTGTCAAGTGTGATGTGACGGCGCGTACGA 2219  
Db 2210 CACCATCCCCAGACAGGCTCAGGAGGTGATCGCAGCATATCAAAACCCAGAACAC 2269  
QY 2220 CACCATCCCCAGACAGGCTCAGGAGGTGATCGCAGCATATCAAAACCCAGAACAC 2279  
Db 2270 GTACTGCTGCTGCGTATGCGGTGCTCCAGAGGCGCCCATGGCACGTCGCCAAGGC 2329  
QY 2280 GTACTGCTGCTGCGTATGCGGTGCTCCAGAGGCGCCCATGGCACGTCGCCAAGGC 2339  
Db 2330 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGC 2389  
QY 2340 CTTCAAGAGCCACGCTCTACTTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGC 2399  
Db 2390 TCACCTGAGGAGACACGCGCTGAGGGATGCGGTGCTATCGACGACAGCTCCTCCCT 2449  
QY 2400 TCACCTGAGGAGACACGCGCTGAGGGATGCGGTGCTATCGACGACAGCTCCTCCCT 2459  
Db 2450 GAATGAGCCACGAGTGGCCTTTCGAGCTTCTTCTAGCTTTCATGTCACACGCGGT 2509  
QY 2460 GAATGAGCCACGAGTGGCCTTTCGAGCTTCTTCTAGCTTTCATGTCACACGCGGT 2519  
Db 2510 GCGCATCAGGGGCAAGTCTCTACGTCAGTGCAGGGGATCCCGAGGCTCCATCCTCTC 2569  
QY 2520 GCGCATCAGGGGCAAGTCTCTACGTCAGTGCAGGGGATCCCGAGGCTCCATCCTCTC 2579  
Db 2570 CACGCTGCTCTGACGCTGTGCTACGCGGACATGGAGAACAGCTGTTTTCGGGGATTCG 2629  
QY 2580 CACGCTGCTCTGACGCTGTGCTACGCGGACATGGAGAACAGCTGTTTTCGGGGATTCG 2639  
Db 2630 CGGGACGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2689  
QY 2640 CGGGACGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2699  
Db 2690 CCAGCGAAAAACCTTCTCTCAGSACCTGCTCCGAGGCTGCTCCTGAGTATGCTCGGTGT 2749  
QY 2700 CCAGCGAAAAACCTTCTCTCAGSACCTGCTCCGAGGCTGCTCCTGAGTATGCTCGGTGT 2759

Db 2750 GAACCTCCGGAACACAGCTGTGACCTTCCCTGTAGAAGACGAGGCCCTGGTGCCACGGC-2809  
QY 2760 GAACCTCCGGAACACAGCTGTGACCTTCCCTGTAGAAGACGAGGCCCTGGTGCCACGGC 2819  
Db 2810 T 2810  
QY 2820 T 2820  
RESULT 3  
LOCUS AF043739 8960 bp DNA SYN 05-APR-1998  
DEFINITION Synthetic construct human telomerase catalytic subunit  
(hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete  
sequence.  
ACCESSION AF043739  
NID AF023054  
KEYWORDS synthetic construct.  
SOURCE synthetic construct.  
ORGANISM artificial sequence.  
REFERENCE 1 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D.,  
Haber,D.A. and Weinberg,R.A.  
TITLE Telomerase activity is restored in human cells by ectopic  
expression of hTERT (hEST2), the catalytic subunit of telomerase  
Oncogene 16 (9), 1217-1222 (1998)  
JOURNAL 9818713  
MEDLINE  
REFERENCE 2 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D.,  
Haber,D. and Weinberg,R.A.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center,  
Cambridge, MA 02142, USA  
FEATURES  
Location/Qualifiers  
source 1..8960  
/organism="synthetic construct"  
/db\_xref="taxon:32830"  
order(1..1101,4606..8960)  
/note="mammalian expression vector; see GenBank Accession  
Number U47120"  
/organism="Cloning vector pCI-neo"  
/db\_xref="taxon:45772"  
1102..4542  
/note="telomerase catalytic subunit hTERT/hEST2 mRNA"  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.33"  
1102..4542  
/gene="hTERT/hEST2"  
1147..4605  
/note="telomerase/reverse transcriptase"  
/codon\_start=1  
/transl\_table=11  
/product="telomerase catalytic subunit"  
/db\_xref="PID:g3023055"  
/translation="MPRAPRCRAVRSLLRSHYREVLPATFVRRLPGQGRRLVQRGDP  
AAFRALVACLVCPNDARPPPAAPSFQVSCLELYVARVLQRLCERGAKNVLAFOGA  
LLDARGGPEAFTTSYRSLPNTVDALRGSGAWGLLLRRVGDVLLVLLARCALFV  
LVAPSCAYQVCGPLYGATQAPPPHAGSPRRIGRCERAWNHSVREAGVPLGPA  
PGARRGGSASRLPLKRPARGAAPPEPTVQSGSWHPGTRGTPSDRGFCVSPA  
RPAEATSLGALSGRHSHPSVQROHAGPPTSRPPRMDTPCPVVAETKHLYS  
SGDEOLRPSFLSSLRISITGARRLVETFLGSRWPMCTPRRLPQRYQWMPRL  
FLELLNHAQCPYGVLLKTHCPRLAAVTAAGVCAREKPGQSVAAAREEDTDPRRLVQ  
LLQHSSPMQVYGFVACLRLLVPPGLMGRNHRRLNRTKKFISLGRHAKLSLQEL  
TKMSVRGCAWLRSPGCVPAAEHRLREEILAKFLHMLSVYVVELLSRFFVTET  
TFQKRLFFYKRSWKLQSIGIRQHLKRVOLRELSAEVROHREARPAALLTSRLRFI  
PKDGLPIVNDYVVGARFREREKRAERLTSVKALFVNLTERARPGLLGASVIG  
LDDIHRAMRTVLRRAQDPPPELIFVKVDVTGAYDIPQDRLTEVIASTIKPQNTYC  
VRYAVVQAAHGHVRAKFSHSTLTDLPYMRQFVAHQETSPLEDAVVIIPQSSSL  
NEASSGLFVDFRMCCHAVIRKGSVQCGIPQGSILSLCSCYGDGMENKLFAG  
IRDDGLLLRVDLFLVTHLTHAKTFLRTLVRGVEYGVNLRKTVVNFVEDEAL  
GGTAFVQMPAHGFLPWCGLLDLDTLEVDOSDSYXARTSIRASLTNRGCFKAGNRWR  
KLFGLVRLKCHSLFLDLQVNSLQVCTNIYKILLQNYRHFACVQLQPHQQVWKAPT  
FFURVLSDTASLFLSILKAKNAGMSIGKAAGALPSEAVQWICHQAFLLKLRRHRT  
YVPLGLSRLTAQOLSKRLPGTLTLTALEAANPALPSDFKTLID"

GGTAFAVQMPAHGFLPWCGLLLDTRILEVQSDYSYARTSIRASITFNRGFKAGRNMR  
KLFGLRLKCHSLFDLQVNSLOVCTNIYKILLQAYREHACVLOLPFFHOQYKNP  
FFLRVSDTSLCYSILKAKNAGSLGAKAGPLSEAVQWILCHQFLKLRHRT  
YPLUGSLRATQOLSKRLPGLTITALEAANPALPSDFKILTDIPYDVPYAGPYD  
VPDYAG"

source 4543 .4605

/note="engineered based on the epitope for the  
anti-influenza virus hemagglutinin monoclonal antibody  
12CA5; similar to part of the Influenza A hemagglutinin  
gene sequence in GenBank Accession Number M29257"  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others  
ORIGIN

Query Match 100.0%; Score 901; DB 31; Length 8960;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3000 GTCCAGACTCGCTTCATCCCAAGCCGACGGCTGCGGGCGATTGTGAACATGGACTA 3059  
|||||  
Qy 1920 GTCCAGACTCGCTTCATCCCAAGCCGACGGCTGCGGGCGATTGTGAACATGGACTA 1979  
|||||  
Db 3060 GTCGTGGGAGCCAGACGTTCCGAGAGAAAAGAGGCGGAGGCTCACTCGAGGCT 3119  
|||||  
Qy 1980 GTCGTGGGAGCCAGACGTTCCGAGAGAAAAGAGGCGGAGGCTCACTCGAGGCT 2039  
|||||  
Db 3120 GAAGCACTGTTACGCGTCACTACGAGGCGGCGGCGGCGGCGGCTCCCTGGGCGC 3179  
|||||  
Qy 2040 GAAGCACTGTTACGCGTCACTACGAGGCGGCGGCGGCGGCGGCTCCCTGGGCGC 2099  
|||||  
Db 3180 CTCGTGCTGGGCGTGAGAGATATCCAGAGGCGGCGGCGGCGGCTTCGTGCTGCGG 3239  
|||||  
Qy 2100 CTCGTGCTGGGCGTGAGAGATATCCAGAGGCGGCGGCGGCGGCTTCGTGCTGCGG 2159  
|||||  
Db 3240 GGCCAGGACCGCGCTGAGCTACTTGTCAAGTGGATGACGGCGGCGGTACGA 3299  
|||||  
Qy 2160 GGCCAGGACCGCGCTGAGCTACTTGTCAAGTGGATGACGGCGGCGGTACGA 2219  
|||||  
Db 3300 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCCAAGAAC 3359  
|||||  
Qy 2220 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCCAAGAAC 2279  
|||||  
Db 3360 GTACTGCGTGGTGGTATGCGGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 3419  
|||||  
Qy 2280 GTACTGCGTGGTGGTATGCGGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 2339  
|||||  
Db 3420 CTTCAAGAGCCAGCTCTACTCTGACAGACCTCCAGCGGTACATGCGCAGAGTTCGTGCG 3479  
|||||  
Qy 2340 CTTCAAGAGCCAGCTCTACTCTGACAGACCTCCAGCGGTACATGCGCAGAGTTCGTGCG 2399  
|||||  
Db 3480 TCACCTGAGGAGACCCGCGCTGAGGAGTCCGCTGCTCATCGAGCAGAGCTCCCTCCCT 3539  
|||||  
Qy 2400 TCACCTGAGGAGACCCGCGCTGAGGAGTCCGCTGCTCATCGAGCAGAGCTCCCTCCCT 2459  
|||||  
Db 3540 GAATGAGGCGCAGCTGGCTCTTCGAGCTCTTCCTAGCTTCATGTGCCACAGCGCGT 3599  
|||||  
Qy 2460 GAATGAGGCGCAGCTGGCTCTTCGAGCTCTTCCTAGCTTCATGTGCCACAGCGCGT 2519  
|||||  
Db 3600 GCGCATCAGGCGCAGCTCTTCGCTCAGTCCAGGCGGATCCGCGAGGCTCCATCCTCTC 3659  
|||||  
Qy 2520 GCGCATCAGGCGCAGCTCTTCGCTCAGTCCAGGCGGATCCGCGAGGCTCCATCCTCTC 2579  
|||||  
Db 3660 CACGCTGCTGTCAGCCTGCTGTCAGGCGGACATGAGAACAGCTGTTGCGGGGATTTCG 3719  
|||||  
Qy 2580 CACGCTGCTGTCAGCCTGCTGTCAGGCGGACATGAGAACAGCTGTTGCGGGGATTTCG 2639  
|||||  
Db 3720 GCGGACCGGCTGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3779  
|||||  
Qy 2640 GCGGACCGGCTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2699  
|||||  
Db 3780 CCACGCAAAACCTTCCTCAGACCGCTGCTCGAGGCTGCTCGTGTGATGCTGCTGCTGCT 3839  
|||||

Qy 2700 CCACGCGAAAACCTCTCCTCAGGACCCCTCGAGGCTGCTCCCTGAGTATGGCTCGGTGGT 2759  
Db 3840 GAACCTCGGAAGACACTGGTGAACCTCCCTGTAGAGACGAGGCGCTCGGTGCGACGCGC 3899  
|||||  
Qy 2760 GAACCTCGGAAGACACTGGTGAACCTCCCTGTAGAGACGAGGCGCTCGGTGCGACGCGC 2819  
|||||  
Db 3900 T 3900  
Qy 2820 T 2820  
RESULT 4  
LOCUS AF073311 3369 bp mRNA ROD 09-SEP-1998  
DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.  
ACCESSION AF073311  
NID g3551846  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3369)  
AUTHORS Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.  
TITLE Expression of mouse telomerase catalytic subunit in embryos and  
adult tissues  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)  
MEDLINE 98393668  
REFERENCE 2 (bases 1 to 3369)  
AUTHORS Martin-Rivera, L., Herrera, E. and Blasco, M. A.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of  
Biotechnology, Cantoblanco, Madrid 28049, Spain  
FEATURES  
source  
1..3369  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..3369  
/codon\_start=1  
/product="telomerase catalytic subunit"  
/translation="MTRAPRCAPVRSLLRSRYREVWPLATFVRRLQPEGRRLVQPGDP  
KIYRTVAOCLVCMHWSQPPADLSFHOVSSILKELVARVORLCERNERNVLAQFE  
LINEARGPPMAFTSSVRSYLPNTVETLTVSGAWMLLSRNGVDDLVLVLAHLCALYL  
LVPSCAYQVCGSLFQICATDIFSVSASIRTPRPVVEGHPHQVLPFTSGKSWSPF  
APPLALPGRGTRKHLSTSVPSAKKACYPVPRVEGHPHQVLPFTSGKSWSPF  
ARSPEYPTAEKDLSSKGVSDLSGKVCCKKPSSTLSLSPRQAFQLRPLFTRH  
FLYSGDQERLNPSSLNLPNLGARRLVEIIFLSRPTSGPLRHLRSRYW  
OMRPLFQILLVNAECQYVRLRSHCRFTANQVTDALNTSPHMDLLRLHSSPQ  
VYGLFACLCVKVVSASLWTRHNRERFFNLKKEFISLGKGLSLQELMMKMKVECH  
WLKSSPGKDRVPAAEHRLERILATFLWMDTVVOLLRSFFYITESTFQKRLFFY  
RKSWSKLQSIGVROHLRVLRELSEEVRRHODTWLAMPICRLRIFPKNGLRPIV  
NMSYSGTRALGRKQAHFTQRLKTLFSLNLYERTKHPHLMGSSVLGMDIYRTWA  
FVLVRALDQTPRMVFKADVTGAYDAIPQGLVEVVANNIRHSESTYCIROYAVVRR  
DSOGVHKSPROVTTLSLOPYMGOFKLHODSDASALNSVVIQSIISMNSSL  
DFDLHFLRHSVVKIGDRYTCQGIPOGSSSLTLCSCFCGDMENKFLAEVORDGL  
LRVDDFLVTTPLHLDQTLFSLVHGVPEYGCMLNQLKTVNFPVPEPTGLGAAPYQ  
LPKCHFLPWCGLLDTLEFCDYGYAQTSIKTSITLTSQSFQKAGKRNKLLFIIS  
LKCHGLFLQVNSLQVTCINIKYIFLQAYRHFACVIOPLDQVRKRNKLLFIIS  
SQACCCVAILKVKNPGLTKAGSFPPEAAHLCYQAFLLKLAHSHVIYKCLLGLPLT  
AQKLCRLKPEATMTILKAADPALSTDFQTLTD"

BASE COUNT 737 a 959 c 891 g 782 t  
ORIGIN

Query Match

Best Local Similarity 41.0%; Score 369; DB 28; Length 3369;

Matches 667; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

Db 1827 CAGACTCGCTTCATCCCAAGCCCAAGCGCTCGCGGCCCATGTGTAACATGAGTTATAG 1886  
|||||

Qy 1923 CAGACTCGCTTCATCCCAAGCCCAAGCGCTCGCGGCCCATGTGTAACATGAGTTATAG 1982  
|||||

Db 1887 CATGGTACCAAGACTTTGGGCAGAGGAGCCAGCATTTTCAACCCAGCGCTCTCAA 1946  
|||||

QY 1983 CGTGGGACGACGCTTCCGACAGAAAGAGGGCGGCGCTCTCACCTCGAGGGTGAA 2042  
Db 1947 GACTCTCTTCCAGCATGCTCAACTATGAGCGGACAAAACATCTCTACCTTATGGGGTCTTC 2006  
QY 2043 GGCAGCTTCTAGCGTCTCAACTACGAGCGGGCGCGCCCGCTCTCTGGCGGCTC 2102  
Db 2007 TGTAAGCTGATGAATACATCTACAGGACCTGCGGGGCTTGTGCTGGTGGCTGCTGC 2066  
QY 2103 TGTCTGGGCTGACGATATCCAGGGGCTGGCGACCTTGTGCTGGTGGCTGGCGG 2162  
Db 2067 TCTGACACGACACCCAGGATGCTTGTGTAAGGACAGATGTGACCGGGGCTATGATGC 2126  
QY 2163 CCAGGACCCCGGCTGAGCTGTACTTGTCAAGGTGATGAGCGGCGGTACGACAC 2222  
Db 2127 CATCCCCAGGTAAGCTGTGGAGTGTGTGCAATATGATCAGGCACTCGGAGACAC 2186  
QY 2223 CATCCCCAGGACAGGCTACGAGGCTCATCGGAGCATCATCAACCCC---AGAACAC 2279  
Db 2187 GTACTGTATCGGACATGACAGTGGTCCGAGAGATAGCAAGGCCAAGTCCACAGTC 2246  
QY 2280 GTACTGGCTGCGTATGCGGTGGTCCAGAGGCGCGCCATGGGACACGCTCGCAAGGC 2339  
Db 2247 CTTTAGGACAGATGACACACCTCTCTGACCTCCAGCCATATGAGCGGCTTCTCTTAA 2306  
QY 2340 CTTCAAGAGCAGCTCTTACCTTGACAGACCTCCAGCGGTACATGCGAGTGTCTGGC 2399  
Db 2307 GCATCTCAGGATTCAGATGCCAGTGCACGTGAGGAATCCGTTGTATCAGCAGCAGCAT 2366  
QY 2400 TCACCTCAGAGAC-----CAGCCCGCTGAGGATGCGCTGTCATCGAGCAGAGCTC 2453  
Db 2367 CTCTATGAATGAGAGCAGCAGCAGCTTGTGATCTTCTTCTGACATCTCTGCTGACAG 2426  
QY 2454 CTCCCTGAATGAGCCAGCAGTGGCTCTTCCAGAGCTTCTTACGCTTCTATGTGCCACCA 2513  
Db 2427 TGTCTAAAGATTTGTCAGAGTGTCTATGACAGTGCAGGCGCATCCCGAGGCTCCAG 2486  
QY 2514 CGCGTCCGATCAGGGGCAAGTCTTACGTTTGTGATGACCTTCTGTTGTGAGCGCTCA 2573  
Db 2487 CCTATCACCCCTGCTGTCAGTCTGTGTTTCCGAGACATGAGAACAAAGCTGTTGTCTGA 2546  
QY 2574 CCTCTCCAGCTGCTGTCAGCCTGTGTCAGGCGGATGAGAACAAAGCTGTTGTGCGG 2633  
Db 2547 GGTCCAGCGGATGGTGTCTTACGTTTGTGATGACCTTCTGTTGTGAGCGCTCA 2606  
QY 2634 GATTCCGCGGAGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2693  
Db 2607 CTTGACCAAGCAAAACCTTCTCAGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2666  
QY 2694 CTTACCCACCGAAGACCTTCTCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2753  
Db 2667 CATGATAAAGTTCAGAGACAGTGGTGAATCTTCCCTGTGAGGCTGCTGCTGCTGCTGCTG 2726  
QY 2754 CGTGGTGAATTCGGAAGACAGTGGTGAATCTTCCCTGTAGAACAGGAGGCGCTGCTGCTG 2813  
Db 2727 TGCAGCT 2733  
QY 2814 CACGGCT 2820

RESULT 5  
LOCUS AF051911 3426 bp mRNA ROD 02-APR-1998  
DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.  
ACCESSION AF051911  
NID g3005591  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3426)  
AUTHORS Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and DePinho,R.A.  
TITLE Expression of mouse telomerase reverse transcriptase during development, differentiation, and proliferation

JOURNAL Oncogene (1998) In press  
REFERENCE 2 (bases 1 to 3426)  
AUTHORS Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and DePinho,R.A.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-1998) Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA

FEATURES  
source Location/Qualifiers  
1..3426  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="13"  
/map="MGD offset position 40; 40 cm from centromere"  
30..3398  
/note="mTERT"  
/codon\_start=1  
/product="telomerase reverse transcriptase"  
/db\_xref="PID:g3005592"  
/translation="MTRAPRCFAVRLSRREVWPLATFVRRLGPEGRRLVQPDPP  
KIYRLVAQCLVCMHWSQPPADLSFHQVSSKLKELVARVQRLKERNRNVLAFQF  
LNEARGPPMAFTSVRSYLPNTVETLRVSGAWMLLSRVDLLVYLALCALYL  
LVPPCAYQVCGSPLYQICATTDIWPVSASRYRTPVGRNFTNRLFLQIKSSROE  
APKPLALPSRGTKRHLSTSTVPSAKKACYPVPRVREGPHROVLPTSGKSWPSP  
ARSPVPTAEKDLSSGKVSDLSGSCVCKHKPSTSLSPRQNAFOLRPIETRH  
FLYSRGDSGLNPSFLLSNQLPNTGARRLVEIIFLGSRPRTSGPLCTHRLSRRIW  
QMRPLFQQLLVNHAECQYVRLLRSHCFRTANOQVTDALNTSPHMLDLRLHSSPWQ  
VYGFRLACLVVYSASLWGRNRRERFKLKKFISLGRYKLSQLSDELMMKRWKVEDCH  
WLRSSPKDRVPAEAHRLERILATFLWMDTVVOLLRSFYITESTFKNRLFFY  
RKSWSKLQSIGVROHLERLRELSEQEVRRHODTWMAMPICRIRFIKPNGLRITV  
NMSYSGTFRALGRKQAQHTORLKTLSMLNYERTKHPHMGSSVLGMDNIYRTWRA  
FVIRVALDQTPMTFVKADVTGADYDAIPQGLKEVYVANNIRHSESTYCIQIYAVRR  
DSQGVHKSFRRVTTLSDLQPMGOFKHLQSDASALRNSVTEQSTMSHRESSLL  
FDFFHLFHSVYKIGDRCTQCGIPQGSLSLTLCSLCFDMENKLFQVQDGLL  
LRFVDDFLVTLPHLQAKTFLSLVHGVPEYGMINLQKTVNPEVPTGLGAAPYQ  
LPAHCLFPMCGLLDTQLEVEDKYSYQAQTSIKTSLTFSQSVFKAGKTWRNKLVLRL  
LACHGLFDLQVNSLQTVCLINIKYIFLQAYRFHACVIOLPFDQVRKNLTFLGLIIS  
SQASCCYAILKVNFGMTLKSGSFPPEAAHMLCTYQAFLLKLAHNSVIYKCLLGLPIIT  
AOKLLCRKLPEATMTILKAAADPALSTDFQTILD"

BASE COUNT 746 a 979 c 907 g 794 t

Query Match 41.0%; Score 369; DB 28; Length 3426;  
Best Local Similarity 73.5%; Pred. No. 7.36e-230;  
Matches 667; Conservative 0; Mismatches 291; Indels 9; Gaps 2;

Db 1856 CAGACTGCGCTTCATCCCAAGCCCAAGCGCTCGCGCCCATTTGTAACATGAGTATAG 1915  
QY 1923 CAGACTCGGCTTCATCCCAAGCCTGACGGCTCGCGCCGATTTGAACATGAGTACGT 1982  
Db 1916 CATGGGTACAGAGCTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1975  
QY 1983 CGTGGGAGCCAGACGTTCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042  
Db 1976 GACTCTCTCAGCATGCTCAACTATGAGCGGACAAACATCTCTACCTTATGGGGTCTTC 2035  
QY 2043 GGCAGTGTTCAGGCTGCTCAACTACGAGCGGCGCGCGCGCGCGCTCTGGCGCGCTC 2102  
Db 2036 TGTACTGGGTATGAATGACATCTACAGGACCTGCGGGCGCTTGTGCTGCTGCTGCTGC 2095  
QY 2103 TGTCTGGGCTTGAGATATCCAGGCGCTGCGGACCTTCTGCTGCTGCTGCTGCTGCTG 2162  
Db 2096 TCTGGACACAGACCCAGGATGCTTGTGTAAGGACAGATGTGACCGGGCGCTATGATGC 2155  
QY 2163 CCAGGACCCCGCTGAGCTGTACTTTGTCAAGGTGATGTGACGGCGCGGTACGACAC 2222  
Db 2156 CATCCCCAGGTAAGCTGTGAGGTTGTGCAATATGATCAGGCACTCGGAGACAC 2215  
QY 2223 CATCCCCAGGACAGGCTCAGGAGGCTCATCGGAGCATCATCAACCCC---AGAACAC 2279  
Db 2216 GTACTGTATCCGACGATCATGCTGCTGAGGTTGTGCAATATGATCAGGCACTCGGAGATC 2275  
QY 2280 GTACTGCGCTGCGGTATGCCGTGGTCCAGAGAGGCGCGCCATGGGCAAGGC 2339



```

repeat_region 4366. 4661
/rpt_family="Alu"
repeat_region 5327. 5602
/rpt_family="Alu"
repeat_region 6586. 6956
/rpt_family="L1"
repeat_region 6647. 6684
/note="(CA)19"
/rpt_type-tandem
repeat_region 7113. 7373
/rpt_unit-CA
/rpt_family="Alu"
repeat_region complement(7830. 8185)
/standard_name="possible repeat"
repeat_region 8258. 8503
/rpt_family="Alu"
repeat_region 9070. 9387
/rpt_family="Alu"
repeat_region complement(9740. 9845)
/rpt_family="MER42"
repeat_region complement(10440. 11015)
/rpt_family="Alu"
repeat_region 11950. 12250
/rpt_family="Alu"
repeat_region 12057. 12085
/note="(A)29"
/rpt_type-tandem
repeat_region 12365. 12645
/rpt_unit-A
repeat_region 13727. 13750
/note="(AC)12"
/rpt_type-tandem
repeat_region 13783. 14024
/rpt_unit-AC
repeat_region 14175. 14470
/rpt_family="L1"
repeat_region 14175. 14470
/rpt_family="Alu"
repeat_region complement(14906. 15259)
/standard_name="possible repeat"
repeat_region 15300. 15613
/rpt_family="Alu"
repeat_region 16671. 16690
/note="(T)20"
/rpt_type-tandem
repeat_region complement(16675. 16977)
/rpt_unit-T
repeat_region 16993. 17085)
/rpt_family="MER42"
repeat_region complement(17678. 18276)
/rpt_family="Alu"
repeat_region 19305. 19583
/rpt_family="Alu"
repeat_region 19914. 19945
/note="(GTTT)8"
/rpt_type-tandem
repeat_region complement(19943. 20222)
/rpt_unit-GTTT
repeat_region 21202. 21496
/rpt_family="Alu"
repeat_region 21736. 22035
/rpt_family="Alu"
repeat_region 22017. 22038
/note="(A)22"
/rpt_type-tandem
repeat_region 22321. 22457
/rpt_unit-A
repeat_region 22321. 22457
/rpt_family="Alu"
repeat_region complement(22638. 22981)
/standard_name="possible repeat"
repeat_region 23022. 23326

```

---

```

/rpt_family="Alu"
repeat_region 23473. 23761
/rpt_family="Alu"
repeat_region 23744. 23767
/note="(A)24"
/rpt_type-tandem
/rpt_unit-A
repeat_region complement(24639. 24694)
/rpt_family="MER42"
repeat_region complement(25349. 25713)
/rpt_family="Alu"
repeat_region complement(25727. 26471)
/rpt_family="Alu"
repeat_region 27191. 27477
/rpt_family="Alu"
repeat_region 27774. 28057
/rpt_family="Alu"
repeat_region 28040. 28066
/note="(A)27"
/rpt_type-tandem
repeat_region 28712. 28930
/rpt_unit-A
repeat_region complement(28769. 28838)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 28987. 29214
/rpt_family="Alu"
repeat_region join(29485. 29559, 29598. 29779)
/note="95% identity dbSTS:G14522 (SHCC-11312)"
repeat_region 29495. 29976
/standard_name="AI027942"
/misc_feature complement(30401. 30536)
/note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(30682. 30733)
/rpt_family="MIR"
/misc_feature complement(31573. 31724)
/note="GRAIL 2 excellent exon, frame 1"
/misc_feature complement(32159. 32232)
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 32388. 32488
/rpt_family="MLT1"
repeat_region 32617. 32908
/rpt_family="Alu"
repeat_region 32977. 33088
/rpt_family="MLT1"
repeat_region complement(33670. 33785)
/rpt_family="Alu"
/misc_feature complement(34021. 34144)
/note="GRAIL 2 excellent exon, frame 2"
/misc_feature complement(35238. 35331)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 36392. 36663
/rpt_family="Alu"
/misc_feature 36901. 37222
/note="100% identity EST ou55c09.xl"
/note="db xref=dbEST:AI025011"
36901. 37164
/standard_name="STSC-9983"
/db_xref="dbSTS:G26554"
/misc_feature complement(37595. 37654)
/note="GRAIL 2 excellent exon, frame 0"
/misc_feature complement(join(38063. 38218, 38462. 38578, 38741. 38995,
39071. 39203, 39532. 39630, 39935. 40048, 40300. 40410,
40503. 40661, 41868. 41972, 42103. 42225, 42492. 42569,
44379. 44507))
/standard_name="histidyl-tRNA synthetase"
/misc_feature complement(38069. 38215)
/note="78%-100% protein identity GenPept:U18937"
/misc_feature complement(38462. 38578)
/note="GRAIL 2 excellent exon, frame 0"

```

... remainder of annotations omitted.

Query Match 3.8%; Score 34; DB 26; Length 74371;  
Best Local Similarity 15.5%; Pred. NO. 3.54e-06;  
Matches 13; Conservative 46; Mismatches 25; Indels 0; Gaps 0;

Db 15918 KCKCTSTKSCMSKSRGKRGYSWKYRCAMWMTCKSSKWCWSYEMRMKCYSCSYC 15977  
Cp 2609 TCGCGTAGCAGGCTGAGAGCGAGCGTGAGAGATGGACCTCGGGGATCCCTGG 2550

Db 15978 YCSSGKKYWCRCSMYWTYYSYK 16001  
Cp 2549 CACTGGACGTAGACTGCCCTG 2526

RESULT 8  
LOCUS 128278 215 bp DNA PAT 30-OCT-1996  
DEFINITION Sequence 5 from patent US 5569830.  
ACCESSION 128278  
NID 91819054  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 215)  
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.  
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease  
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;  
FEATURES Location/Qualifiers  
source 1..215  
/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others  
ORIGIN

Query Match 3.3%; Score 30; DB 21; Length 215;  
Best Local Similarity 14.8%; Pred. NO. 1.16e-03;  
Matches 16; Conservative 46; Mismatches 46; Indels 0; Gaps 0;

Db 1 MTNVTSSSVSRVASCNDKAKDGNTSSWTTDCNRTNVCBDTITYRVNDSGHNK 60  
QY 2107 CTGGCGCTGAGATATCCACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2166

Db 61 YSSANYGNNVGAHYYHTHTVNSGADSKTVDSYNASTSSNG 108  
QY 2167 GACCCCGCCCTGAGCTGACTTCTGTCTGTCAAGGTGGATGTGACGGGCGC 2214

RESULT 9  
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998  
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.  
ACCESSION AC005369  
NID 93367505  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 74371)  
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.  
TITLE Sequencing of human chromosome 5  
JOURNAL Unpublished  
AUTHORS Rike,D.O.  
REFERENCE 2 (bases 1 to 74371)  
AUTHORS Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System  
JOURNAL Unpublished  
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.

TITLE Direct Submission  
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.  
COMMENT Sequence submitted by:  
DOE Joint Genome Institute.  
FEATURES Location/Qualifiers  
source 1..74371  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="5g"  
/clone="119j3"  
/chromosome="5"  
/note="LBNL H175"  
893..1030  
/rpt\_family="Alu"  
2255..2438  
/rpt\_family="Alu"  
2818..2859  
/note="(GT)21"  
/rpt\_type=tandem  
/rpt\_unit=GT  
join(3246..3410,3721..3828)  
/standard\_name="RUF"  
/note="65% & 69% protein identity GenPept:U22377"  
3431..3724  
/rpt\_family="Alu"  
3707..3728  
/note="(A)22"  
/rpt\_type=tandem  
/rpt\_unit=A  
4366..4661  
/rpt\_family="Alu"  
5327..5602  
/rpt\_family="Alu"  
6586..6956  
/rpt\_family="L1"  
6647..6684  
/note="(CA)19"  
/rpt\_type=tandem  
/rpt\_unit=CA  
7113..7373  
/rpt\_family="Alu"  
complement(7830..8185)  
/standard\_name="possible repeat"  
8258..8503  
/rpt\_family="Alu"  
9070..9387  
/rpt\_family="Alu"  
complement(9740..9845)  
/rpt\_family="MER42"  
complement(10440..11015)  
/rpt\_family="Alu"  
11950..12250  
/rpt\_family="Alu"  
12057..12085  
/note="(A)29"  
/rpt\_type=tandem  
/rpt\_unit=A  
12365..12645  
/rpt\_family="Alu"  
13727..13750  
/note="(AC)12"  
/rpt\_type=tandem  
/rpt\_unit=AC  
13783..14024  
/rpt\_family="L1"  
14175..14470  
/rpt\_family="Alu"  
complement(14906..15259)  
/standard\_name="possible repeat"  
15300..15613  
/rpt\_family="Alu"

[illegible]







gene	GNMPSERKGGGFENKTAQIVADELKVLRYYWLTQSGFVNRVNTLTGCLDILIGTS GDIVQANPYFRSAYIVLARGKEADILKRDDPRLKDRQIGIAGTTPPSNRUSLEKL GERIYATAPYRAGERRKHQVAAEVIADLAEKKIDVAILWGPAAAGWLAKQSGVPMDV PLLHEGRRALTFRVSGVRHNENDMRSLNTVLRRKADIEAVLEVEYVPLLAEDT KPLDAADE"
gene	complement(10966. .11442) /gene="foia"
CDS	complement(10966. .11442) /gene="foia" /function="foliate biosynthesis" /codon_start=1 /transl_table=11 /product="6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase" /db_xref="PID:g2160893"
gene	translocation="MTRAYLGLGSGNIGDKAAMLAGAVEHLAATPGINVARSDYRTP PWGTDQWELNAAVAIDTELTHGLEVLCSIEAALGRVREPRWGPRTVIDIDLVALE GAQVSDERLVLPHFRFERAFVLVPLAEIAPDLVIGGETVREALAKLDPSGIERVE" complement(11437. .11831) /gene="foia"
CDS	complement(11439. .11831) /gene="foia" /function="foliate biosynthesis" /codon_start=1 /transl_table=11 /product="dihydroneopterin aldolase" /db_xref="PID:g2160884"
gene	translocation="MADRILVHRLANVARGVLPPEERLQGRFYISLECKLDLAPAGR SDVVAATSVADLAEIALEIASNRFRFALISALAEIAETCLARFPRIETIAVRIDKPS APIPAVLIDYRAIEIVGRSDSVATTETR" complement(11823. .12618)

...  
Note: remainder of annotations omitted.

Note: remainder of annotations omitted.

	Query Match	3.28;	Score 29;	DB 15;	Length 15807;
	Best Local Similarity 67.9%;	Pred. No. 4.69e-03;			
	Matches 55;	Conservative 0;	Mismatches 26;	Indels 0;	Gaps 0;
Db	3117	CGCCCTGGCCGCCGAACTCTATGCCGGGAGGGGAGCGTCTCGACCTTGACCGACATCCA	31176		
QY	2316	CGCCCATGGGCAGCTCGCAAGGCCTTCAAGAGCCAGCTCTACCTTGACAGACCTCCA	23755		
Db	3177	GCAGATCGTGAGCGAGGTCGT	3197		
QY	2376	GCGGTACATCGCACAGTTGCT	2396		

RESULT	11
LOCUS	128278 215 bp DNA
DEFINITION	Sequence 5 from patent US 5569830.
ACCESSION	128278
NID	G1819054
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 215)
TITLE	Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H. Plant inhibitors of fungal polygalacturonases and their use to control fungal disease
JOURNAL	Patent: US 5569830-A 5 29-OCT-1996;
FEATURES	Location/Qualifiers 1..215 /source="unknown"
BASE COUNT	15 a 8 c 25 g 26 t 141 others
ORIGIN	

```
Query Match      3.1% Score 28; DB 21; Length 215;  
Best Local Similarity 14.1%; Pred. No. 1.95e-02;  
Matches        62; Mismatches 84; Indels    0; Gaps    0;
```

Dbb

```
35 DCNRTWGVCDTDTTYRVNDSDSHKYSYSSANYGGNNVGAATHTYYTHTNVGASDQTV 94  
| | : |: : : : |: : : : : | : : : : ||:::
```

Cp 2506 ACATGAACGCTAGGAAGACGTCGAAGAGCCCATGCTGGCCCTCATTCAGGAGAGAGCTCT 2447

Db 95 TDSYNAGSTSSNGTGDNRSGADSYGSSKTAMTSRNRTGTANNVDSRNMNGDAVSGED 154

Cp 2446 GCTCGATGACGAGCGATCCCTCAGCGGGCTGGTCTCTCGAGGTGACGACGACGACGTC 2387

Db 155 KNTKKHAKNSADGKVGSKNNDRNNRYGTGTSKNVSNNGGNGRRDVSYY 204

Cp 2386 GCATGTCGGCTGAGGCTCTCAAGGTAGAGACGTGGCTCTTCGAAGGCC 2337

RESULT	12	
LOCUS	AF027499	5591 bp DNA BCT 01-MAR-1998
DEFINITION	Azotobacter vinelandii mannuronan C-5-epimerase (algC) gene, partial cds; and AlgX, alginate lyase (algI), AlgI, and AlgV genes, complete cds.	
ACCESSION	AF027499	
NID	G2920564	
KEYWORDS	.	
SOURCE	Azotobacter vinelandii.	
ORGANISM	Azotobacter vinelandii	
	Eubacteria; Proteobacteria; gamma subdivision; Azotobacteraceae; Azotobacter.	
REFERENCE	1 (bases 1 to 5591)	
AUTHORS	Vazquez,R.A., Alvarado,D.A., Guzman,J., Soberon-Chavez,G. and Espin,G.	
TITLE	The Azotobacter vinelandii algXLIVFA operon	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 5591)	
AUTHORS	Vazquez,R.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (30-SEP-1997) Microbiologia Molecular, Inst. de Biotecnologia. UNAM, Av. Universidad s/n Col. Chamilpa, Cuernavaca, Morelos 62210, Mexico	

FEATURES  
source

gene

CDS

gene

505

gene

CDS

/gene= arylb  
/note="AlqL: similar to Pseudomonas aeruginosa AlqI."

```

/codon_start=1
/transl_table=11
/product="algininate lyase"
/db_xref="PID:g2920567"
/translation="MKTKLALSLGLLSLLSGAVHAAEALVPPKGVYAPVDIRKGEA
PACPVPEPTCLVFRSKYEGSDAARSTLNEAEAKFRKTAPIQTQIERGVSEWYMW
YMEKGRAGDLECTLAWDAWEDGALLTEYNTGKSMRWALGSLAGAYLRUKTSFS
QPLAAYPEQARLSEFVAKVDGQIKDSDLPKRNHSHYAAWAAVMAAGVATNRPR
LFEDVAQEYFHIAAGQVDSNGFLPNEIKRRQRALAYHNYSLPPLMVAALFALANGVDR
GNDGALGRLAGNVLAGVEKPEPFAERAGDEDDMEDLETDAKFSWLEPYCALYSCSP
ALURERAEMGPFKNRGLGGDVTRIEDPAEKSPRSTVGKRD"
2897..4396
/gene="algI"
/genes="algI"
/notes="similar to Pseudomonas aeruginosa AlgI"
/codon_start=1
/transl_table=11
/product="AlgI"
/db_xref="PID:g2920568"
/translation="WVFSSNVFLFLPLFLGLYLSPARYNLLLLTASYVYAWVR
IDFLGFAAVTENVYNGIRUGAAGVRTSOAQRWLTLGVVVDLCVLGYFYKAYNGVDS
LNAIITSGFVEPEVTHILLPGISFYVPEISYIIDVYEGDTPATRNLDVFAAFVIL
FPHLIAGVLRERLDVQDNHRHTLDKFAEGCTRFMQGFIKKVFTLADSLAPDACHFC
AUSDPTGDWALGALYTAQLYDFDYSDMAIGLMIGFMFMENFNQPIYSOISTE
FWRWHISLSTWRUDYLTSLGNGRGTTFQTYRNLLITMLLGLLGHGANWTFIIGWAG
HGTWLAIERALRIDAAPKTIIRPLRWVFAFLVMVGVVIFRAENLDVAVRWYAAAFPGI
DWTLSLNKAQLTSLOIATILLAYVIAVYGIQFVAQPLTGAPKAKAKANDOADAPQGI
VHASSGTLAYSOAIDIPALATRVAVLLLPFAASVLIKLSAQSFSPFLYFQF"
4409..5575
/gene="algV"
/genes="algV"
/notes="similar to Pseudomonas aeruginosa AlgJ"
/codon_start=1
/transl_table=11
/product="AlgV"
/db_xref="PID:g2920569"
/translation="MKNRTNLVAGTFAGTLLALLSLWSLGAAGFSTADNTPYLVNKLK
ALAFEHYDEEFPKIKGLTNLVAALDYTLFEGCRPGVVICANOWILSPDEEKKTAAS
QNTEDNQALIQGVREILARNNVQLWMAILPAKARLYPENFGQQFASLHEOLLQNFRR
IVADGQAPDLGLPQOAKAGGVFLRTDHTWTPYGAQVAGOLATHTIKPIGVLPES
GVNYVETLPGGKHGDLTNFLPDLFELLPDPQAKHNTROQFESAPAGDDILFA
ETQVPVALVGTYSADERNWFAGLQALGSDLVNFAEDGRPLLPMLKFLQSEDEFKK
SPRLVWIEFFPERYLPWAYDLSEFDADWIAQLKAAGRQDKQLADNTATNQGRH"
BASE COUNT 1051 a 2011 c 1640 g 889 t
ORIGIN

Query Match 2.9%; Score 26; DB 15; Length 5591;
Best Local Similarity 82.5%; Pred. No. 2.66e-01;
Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 3830 ATCTCACCATGCTGCTCGGGCGCTGTGGCAGCGCGCCA 3869
||||| ||| ||||| || ||||| || ||||| || ||||| ||
QY 2572 ATCTCTCCACGCTGCTCTGCAGCGCTGTGTACGCGGACA 2611

RESULT 13
LOCUS RRU02553 1908 bp DNA ROD 19-OCT-1993
DEFINITION Rattus rattus protein tyrosine phosphatase gene, complete cds.
ACCESSION U02553
NID 9409022
KEYWORDS black rat.
SOURCE Rattus rattus
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1905)
AUTHORS Qian,Z., Gilbert,M.E. and Kandel,E.R.
TITLE Differential induction by neuronal activity in rat brain of a
protein tyrosine phosphatase containing a nuclear-localization
signal

```



**THIS PAGE BLANK (USPTO)**



Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlavet@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

## FEATURES

source  
1. 409  
/organism="Homo sapiens"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"  
/db\_xref="ATCC (inhost):158964"  
/db\_xref="taxon:9606"  
/clone\_lib="Jurkat T-cells VI"  
/cell\_type="T-lymphocyte"  
<1..>409

BASE COUNT 65 a 120 c 133 g 86 t 5 others  
ORIGIN

Query Match 17.9%; Score 161; DB 24; Length 409;  
Best Local Similarity 100.0%; Pred. No. 7.58e-249;  
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GTTGGTGGATGATTTCTTTGGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 60  
QY 2660 GTTGGTGGATGATTTCTTTGGTGACACCTCACCTCACCCACCGGAAACCTTCCTCA 2719

Db 61 GGACCTCGTCCGAGGTGCTCCCTAGTAGTGGCTGGTGGTGAACCTGCGGAGACAGTGG 120  
QY 2720 GGACCTCGTCCGAGGTGCTCCCTAGTAGTGGCTGGTGGTGAACCTGCGGAGACAGTGG 2779

Db 121 TGAACCTCCCTAGACGAGCGCCCTGGTGGCAGCGCT 161  
|||||

QY 2780 TGAACCTCCCTAGACGAGCGCCCTGGTGGCAGCGCT 2820  
|||||

RESULT 2 AA281296 389 bp mRNA EST 14-AUG-1997  
LOCUS  
DEFINITION Zt08g02.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:712562 5',  
mRNA sequence.

ACCESSION AA281296

NID g1924194

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominoidea;

Homo.

1 (bases 1 to 389)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.

Insert Length: 2187 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 385.

Location/Qualifiers

1..389

/organism="Homo sapiens"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-  
3']. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="IMAGE:712562"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
<1..>389

## FEATURES

BASE COUNT 87 a 102 c 123 g 77 t  
ORIGIN

Query Match 17.6%; Score 159; DB 8; Length 389;  
Best Local Similarity 100.0%; Pred. No. 5.81e-245;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 231 GTCCAGACTCGCTTCATCCCAAGCCTGACGGCTCGCGCGATTGTGAACATGGACTA 290  
|||||

QY 1920 GTCCAGACTCGCTTCATCCCAAGCCTGACGGCTCGCGCGATTGTGAACATGGACTA 1979  
|||||

Db 291 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGCGCGCTCTCACCTCGAGGGT 350  
|||||

QY 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGCGCGCTCTCACCTCGAGGGT 2039  
|||||

Db 351 GAAGGCACTGTTAGCTGCTCAACTAGCAGCGCGCGCG 389  
|||||

QY 2040 GAAGGCACTGTTAGCTGCTCAACTAGCAGCGCGCGCG 2078  
|||||

RESULT 3 AA754459 252 bp mRNA EST 20-JAN-1998  
LOCUS  
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa  
cDNA clone 97SN1787, mRNA sequence.

ACCESSION AA754459

NID 92801165

KEYWORDS EST.

SOURCE rice.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;

Poales; Poaceae; Oryza.

1 (bases 1 to 252)

Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

Contact: Eun M.Y.

Department of Cytogenetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyungdo, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeun@sun20.asti.re.kr

Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji

University, Yongin, Korea. 449-728 bnhnm@bioserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

Location/Qualifiers

1..252

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db\_xref="taxon:4530"

/clone="97SN1787"



Db	79	TBYISNNVDNTGCGTGVGKTTNVNHSWNNRCSNVVYWBTAICYDBHYBDRANHVD	138
Qy	2411	AGACGACCGCGCTGAGGGATGCGCTCATCGACGAGAGCTCCTCCTCGAATGAGCCCA	2470
Db	139	CTNDRYCNCTASDNGTSATKRVCTGYDKTSDCGGQWKRVTYGSBYBCGVNVAVRT	198
Qy	2471	GCAGTGGCCTCTTCGACGCTTCTCCACGTTCATGTGCCACGCGCGCATCAGGG	2530
Db	199	TSMTWTKSTKMSMDMSRRSRVHYGRW	225
Qy	2531	GCAAGCTCTACGTCCAGTGCCAGGGA	2557
RESULT	6		
LOCUS	AA754458	247 bp	mRNA
DEFINITION	97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa	EST	20-JAN-1998
ACCESSION	AA754458		
NID	92801164		
KEYWORDS			
SOURCE	rice.		
ORGANISM	Oryza sativa		
REFERENCE	Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;		
AUTHORS	Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza. 1 (bases 1 to 247)		
TITLE	Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P., Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y., Lee,M.C. and Eun,M.Y.		
JOURNAL	Large-scale Sequencing Analysis of ESTs from Rice Immature Seed		
COMMENT	Unpublished (1998)		
FEATURES	Contact: Eun M.Y.		
source	Department of Cytogenetics		
	National Inst. Of Agri. Sci. and Tech, RDA		
	Suwon, Kyunggi-do, Korea		
	Tel: 82 331 290 0301		
	Fax: 82 331 290 0307		
	Email: myeunesun20.asti.re.kr		
	Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji		
	University, Yongin, Korea. 445-728 bhnam@bioserver.myongji.ac.kr		
	Seq primer: M13 Reverse Primer.		
	Location/Qualifiers		
	1. .247		
	/organism="Oryza sativa"		
	/cultivar="Wilyang23"		
	/note="Vector: pBluescript SK(+); Site1: EcoRI; Site2:		
	XhoI; Directional cDNA library inserted into lambda ZAPII		
	vector at 5'end with EcoRI and 3' end with Xho I site."		
	/db_xref="taxon:4530"		
	/clone="97SN1784"		
	/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"		
	/cissue_type="Immature Seed"		
	/dev_stage="5 days after pollination"		
	/lab_host="E. coli SOLR"		
BASE COUNT	7 a	16 c	21 g
ORIGIN	34 t	169 others	
Query Match	4.9%	Score 44;	DB 12; Length 247;
Best Local Similarity	12.7%;	Pred. No. 8.5le-34;	
Matches	14; Conservative	60; Mismatches 36;	Indels 0; Gaps 0;
Db	111	SNSVWVWBTAICYDBHYBDRANHVDTRCTNDRGYCNATSDNGTSATKRVCTGYDKTDS	170
Cp	2133	GGCGCTGTGGATATCTCCAGGCCAGCAGACAGCGCCAGGAGCGCGGCGCGCGC	2074
Db	171	DCGGCGKRVYGSBYBCGVNVAVRTSMWTDKSTKMSMDMSRRSRV	220
Cp	2073	CCGCTCGTAGTCAGCACGCTGACAGTCCTTCACCTCGAGGTGAGAC	2024
RESULT	7		

LOCUS	RA200728	375 bp	mus musculus	cdna	EST	19-FEB-1997		
DEFINITION	mus3H09.1 Soares 2NbMT Mus musculus cdna clone 639329 5', mRNA sequence.							
ACCESSION	AA200728							
NID	g1795738							
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus							
REFERENCE	1 (bases 1 to 375)							
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.							
TITLE	The WashU-HMI Mouse EST Project							
JOURNAL	Unpublished (1996)							
COMMENT	<p>Contact: Marra M/Mouse EST Project  WashU-HMI Mouse EST Project  Washington University School of Medicinep  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  Tel: 314 286 1800  Fax: 314 286 1810  Email: mouseest@watson.wustl.edu  This clone is available royalty-free through LLNL ; contact the  IMAGE Consortium (info@image.llnl.gov) for further information.  MG1:391321  Seq primer: -28M13 rev2 from Amersham  High quality sequence stop: 361.  Location/Qualifiers  1..375  /organism="Mus musculus"  /strain="C57BL/6J"  /notes="Vector: p7R3D-Pac (Pharmacia) with a modified  polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA  was primed with a Not I - oligo(dT) primer [5',  TGTTACCAATCTCAAGTGGGAGCGCGGTTTTTTTTTTTTTTTTTTTT  3']; double-stranded cDNA was ligated to Eco RI adaptors  (Pharmacia), digested with Not I and cloned into the Not I  and Eco RI sites of the modified p7R3 vector. RNA  provided by Dr. Bertrand Jordan. Library went through two  rounds of normalization, and was constructed by Bento  Soares and M.Fatima Bonaldo."  /db_xref="taxon:10090"  /clone="639329"  /clone_lib="Soares 2NbMT"  /sex="male"  /tissue_type="Thymus"  /dev_stage="4 weeks"  /lab_host="DH10B"  &lt;1..&gt;375</p>							
BASE COUNT	80 a	95 c	105 g	95 t				
ORIGIN								
Query Match	3.7%; Score 33; DB 23; Length 375;							
Best Local Similarity	75.4%; Pred. No. 4.00e-17;							
Matches	49; Conservative	0; Mismatches	16; Indels	0; Gaps	0;			
Db	144	AGAGCATCTTCATGAATGAGACGACGACGCGTTTGACTCTCTCGACATCTCGTC 203						
Qy	2447	AGAGCTCGCTCCTCGAATGAGGCCAGACGATGCCCTTCGACGCTTCCTACGCTTCATGT 2506						
Db	204	GTCAC 208						
Qy	2507	GCCAC 2511						
RESULT	8							
LOCUS	AF034173	2275 bp	mRNA	EST	22-DEC-1997			
DEFINITION	Homo sapiens nicon2 contig mRNA, partial sequence, mRNA sequence.							



/note="Vector: Lambda ZAP II (Stratagene); Site1: EcoRI; Site2: XhoI; The mRNA was purified from a cloned population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primered EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene)."  
/db\_xref="taxon:31286"  
/clone\_lib="MVAT4 bloodstream form of serodeme WRATat1.1">343

mRNA	98 a	77 c	86 g	82 t	
BASE COUNT					
ORIGIN					

Query Match 2.9%; Score 26; DB 22; Length 343;  
Best Local Similarity 77.1%; Pred. No. 1.10e-07;  
Matches 37; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

237 CCAGAGCGCCGTCATGAGACATTCGCGACGGCCCTCCGAAGACGACGT 284
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2307 CCAGAAGCGCCGCCATGGCGACGTCGCCGAAGCCCTTCAAGAGCCACGT 2354

```

RESULT 10					
LOCUS	T26788	396 bp	mRNA	EST	06-NOV-1995
DEFINITION	T529 trypanosoma brucei rhodesiense cDNA 5' similar to Ribosomal protein S11 homolog.				
ACCESSION	T26788				
NID	g781238				
KEYWORDS	EST				
SOURCE	African trypanosome primer-T3 primer library-MVAT4 bloodstream form of serodeme WRATat1.1 vector-Lambda ZAP II (Stratagene)				
REFERENCE	Rsitel-EcoRI Rsitel2-XhoI The mRNA was purified from a cloned population of bloodstream trypanosomes reexpressing the MVAT4 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primered EcoRI/XhoI cDNA library was constructed in lambda ZAP II (Stratagene).				
ORGANISM	Trypanosoma brucei rhodesiense Eukaryotae; mitochondrial eukaryotes; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma; Trypanosoma brucei.				
AUTHORS	1 (bases 1 to 396)				
TITLE	El-Sayed,N.M.A.				
JOURNAL	CDNA expressed sequence tags of Trypanosoma brucei rhodesiense provide new insights into the biology of the parasite				
COMMENT	Mol. Biochem. Parasitol. 73, 75-90 (1995)				

Contact: El-Sayed NMA  
John Donelson's Laboratory  
Howard Hughes Medical Institute  
300 EMRB, Dept. of Biochemistry, University of Iowa, Iowa City, IA 52242  
Tel: 3193356918  
Fax: 3193356764  
Email: nelsayed@vaxa.veeg.uiowa.edu.

FEATURES	Location/Qualifiers
source	1..396
	/organism="Trypanosoma brucei rhodesiense"
mRNA	<1..>396
BASE COUNT	103 a 101 c 87 g 92 t 13 others
ORIGIN	

Query Match 2.9%; Score 26; DB 19; Length 396;  
Best Local Similarity 77.1%; Pred. No. 1.10e-07;  
Matches 37; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

89 CCAGAGCGCCGTCATGAGACATTCGCGACGGCCCTCCGAAGACGACGT 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2307 CCAGAAGCGCCGCCATGGCGACGTCGCCGAAGCCCTTCAAGAGCCACGT 2354

```

RESULT 11					
LOCUS	AF034177	660 bp	mRNA	EST	22-DEC-1997
DEFINITION	Homo sapiens ntcon6 contig mRNA, partial sequence, mRNA sequence.				
ACCESSION	AF034177				
NID	q2707739				

[illegible]

```

Db   456 GSAGVTNNHHDWS 468
QY   2037 GGTGAAGGCACTG 2049
| : | : : : : 
RESULT 13
LOCUS      AF034173          2275 bp    mRNA           EST        22-DEC-1997
DEFINITION Homo sapiens ntcon2 contig mRNA, partial sequence, mRNA sequence.
ACCESSION  AF034173
NID         g2707735
KEYWORDS    EST.
SOURCE      human..
ORGANISM    Homo sapiens
            Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2275)
AUTHORS     Tripodis,N. and Ragoussis,J.
TITLE       Generation of a transcription map in the region immediately
            centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
            boundary
JOURNAL      Unpublished
REFERENCE   2 (bases 1 to 2275)
AUTHORS     Tripodis,N. and Ragoussis,J.
TITLE       Direct Submission
JOURNAL      Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
            Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES    source
             1..2275                Location/Qualifiers
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="6"
               /map="6p21.3"
               /clone="ntcon2 contig"
               /tissue_type="liver; brain"
               /dev_stage="fetus"
               /note="similar to Br140"
BASE COUNT  438 a   619 c   470 g   599 t   149 others
ORIGIN
Query Match              2.9%; Score 26; DB 11; Length 2275;
Best Local Similarity 10.0%; Pred.No. 1.10e-07;
Matches                  6; Conservative 37; Mismatches 17; Indels 0; Gaps 0;

Db   1483 RWKRGRKRRNTGMRYRMVMAMMCMAMCMMWYWKMGKKCKWRKYRKYKYSTYY 1542
Cp   2703 GTGGTGAGTGCAGTTGCCACCACAGAAATCATCCACAACCGACGAGCACGCCGTC 2644
| ::::|::: |:::: |::: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
| ::::|::: |:::: |::: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::

RESULT 14
LOCUS      AA848961          301 bp    mRNA           EST        30-APR-1998
DEFINITION EST191723 Normalized rat lung, Bento Soares Rattus sp. CDNA clone
            RUATA4 3' end, mRNA sequence.
ACCESSION  AA848961
NID         g2936501
KEYWORDS    EST.
SOURCE      Rattus sp.
ORGANISM    Rattus sp.
            Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 301)
AUTHORS     Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kervavage,A.R. and Adams,M.D.
TITLE       Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
            Gene Index
JOURNAL      Unpublished (1998)
COMMENT
Contact: Lee, NH
          ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org

```

Search completed: Thu Dec 24 18:28:41 1998  
Job time : 3855 secs.

**THIS PAGE BLANK (USPTO)**

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```
Run on: Thu Dec 24 18:30:54 1998; MasPar time 274.67 seconds
        446.298 Million cell updates/sec
Tabular output not generated.
```

```

Title: >US-08-951-733-13
Description: (1920-2820) from US08951733.seq (2 of 2)
Perfect Score: 901
N.A. Sequence: 1920 GTCCAGACATCCGCTTCTATCC.....AGCCCTGGGTGGCAGCGGT 2820
Comp: CAGGCTGAGGGCGAAGTAG.....TCGGGACCCACCGTGGCGA

```

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0: Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database:

```
n_genesec32
1:part8 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40
```

Statistics: Mean 8.771: Variance 5.114: scale 1.715

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query			ID	Description	Pred. No.
	Score	Match	Length			
1	43	4.8	91	Q51746	Oligonucleotide probe	7.51e-12
c	42	4.7	91	Q51746	Oligonucleotide probe	3.02e-11
c	41	4.6	204	N81164	Base substituted E.co	1.20e-10
	40	4.4	204	N81164	Base substituted E.co	4.75e-10
5	38	4.2	114	Q70465	Generic DNA sequence	7.23e-09
6	37	4.1	114	Q70465	Generic DNA sequence	2.78e-08
7	36	4.0	114	Q70467	Generic DNA sequence	1.06e-07
8	36	4.0	114	Q70468	Generic DNA sequence	1.06e-07
9	36	4.0	114	Q70466	Generic DNA sequence	1.06e-07
c	36	4.0	114	Q70467	Generic DNA sequence	1.06e-07
c	36	4.0	114	Q70470	Generic DNA sequence	1.06e-07
c	35	3.9	114	Q70469	Generic DNA sequence	4.00e-07

```

ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DE 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PA (BECT ) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI; 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 4.7%; Score 42; DB 9; Length 91;
Best Local Similarity 5.6%; Pred. No. 3.02e-11;
Matches 3; Conservative 45; Mismatches 6; Indels 0; Gaps 0;

Db 7 ggcgsvhsyyvvvhhvshhsvvvvhhvshvvhvvhvvhvvhvvhvsv 60
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 2695 GTGAGGTGTACCAACAGAAATCATCCACCAAGCAGGAGCCGCCGCC 2642

RESULT 3
ID N81164 standard; DNA; 204 BP.
AC N81164;
DE 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
PH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 4.6%; Score 41; DB 1; Length 204;
Best Local Similarity 5.6%; Pred. No. 3.02e-11;
Matches 3; Conservative 45; Mismatches 6; Indels 0; Gaps 0;

Db 7 ggcgsvhsyyvvvhhvshhsvvvvhhvshvvhvvhvvhvvhvvhvsv 60
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 2695 GTGAGGTGTACCAACAGAAATCATCCACCAAGCAGGAGCCGCCGCC 2642

RESULT 3
ID N81164 standard; DNA; 204 BP.
AC N81164;
DE 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
PH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 4.6%; Score 41; DB 1; Length 204;
Best Local Similarity 5.6%; Pred. No. 3.02e-11;
Matches 3; Conservative 45; Mismatches 6; Indels 0; Gaps 0;

```

```

Best Local Similarity 18.8%; Pred. No. 1.20e-10;
Matches 27; Conservative 58; Mismatches 59; Indels 0; Gaps 0;

Db 50 aggnycgccgggycgcagcyccgaayycdchvgccgymrtthhrrmbnrvdydnrsd 109
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 2093 AGGAGCGCGGGCGCGCGCGCTCGTAGTTGAGCAGCTGAACAGTCTTCACCCCTC 2034
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110 aaawcyccyrsvkydcynachhdhvyvbbvynvhhnncnccbnhvhcnvbnh 169
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 2033 GAGGTGAGACGCTCGGCCCTCTTTTCTCTGCGAACGTTCTGGCTCCGACGACGTAGTCC 1774
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 170 rnwayvvhddarrddvhccvchccg 193
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1973 ATGTTACATCGCGCGCGCGCGCG 1950

RESULT 4
ID N81164 standard; DNA; 204 BP.
AC N81164;
DE 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
PH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 4.4%; Score 40; DB 1; Length 204;
Best Local Similarity 8.2%; Pred. No. 4.75e-10;
Matches 9; Conservative 59; Mismatches 42; Indels 0; Gaps 0;

Db 80 hvccgymrtthhrrmbnrvdydnrsdaawcyccyrsvkydcynachhdhvyv 139
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2432 CCGTCGTCATCAGCAGAGCTCTCCCTCGTAATGAGCGAGCGCTCTTCGACGTCT 2491
||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 bbvynvhhnncnccbnhvhcnvbnhrrnwayvvhddarrddvhccv 189
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Qy 2492 TCCFAGGCTTCATGTGCCACCCGCCGTCGCGCATCAGGGGCAAGTCCTAC 2541
||| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DE 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;

```

KW	effector domain; concatenated heterofunctional protein; linker;
KW	Direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
FH	Key
FT	misc_feature
FT	Location/Qualifiers
FT	55..60
FT	/*tag= a
FT	/note= "this sequence represents 'Z'; Z can be a
FT	sequence of 6, 9 or 12 nucleotides (see
FT	comments)"
FT	
PN	WO9418318-A:
PD	18-AUG-1994.
PF	01-FEB-1994; UO0977.
PR	01-FEB-1993; US-013416.
PR	30-DEC-1993; US-176500.
PR	31-JAN-1994; US-189331.
PA	{UYNC-} UNIV NORTH CAROLINA.
PI	Fowlkes DM, Kay BK;
DR	WPI: 94-279739/34.
DR	P-PDSB; R65150 and R65151.
PT	Identifying proteins or peptide(s) which bind a ligand - by
PT	screening a recombinant vector library expressing fusion proteins
PT	comprising a binding domain and an effector domain
PS	Disclosure: Page 35: 255pp; English.
CC	Q70455 is a generic DNA sequence used to generate random TSAR (Totally
CC	Synthetic Affinity Reagents) peptides.This generic formula can also be
CC	represented as follows: X(NNB)6(TGC)(NNB)112(NNB)14(TGC)(NNB)3Y. X
CC	and Y are flanking restriction sites (X is not the same as Y) that are
CC	not specified further. Other generic sequences are shown in Q70466-68.
CC	Other specific peptides generated by these generic sequences are shown in
CC	R65151-54. TSARS are concatenated heterofunctional proteins or peptides,
CC	comprising at least two functional regions - a binding domain with
CC	affinity for a ligand and a second effector peptide portion that is
CC	chemically or biologically active.They may further comprise a linker
CC	peptide between the 2 domains.The oligonucleotides are also designed so
CC	that the expressed peptide contains 2 or 4 cysteine residues positioned
CC	in, or flanking, the unpredicted or variant residues. These residues
CC	confer some degree of conformational rigidity to the peptides. The TSARS
CC	or conspans. comprising a TSAR binding domain can be used in vivo to
CC	deliver a chemically or biologically active moiety, eg. metal ion,
CC	radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC	cell. They can also replace the function of macromolecules, eg.
CC	monoclonal or polyclonal antibodies and therefore circumvent the need
CC	for complex methods of hybridoma formation or in vivo antibody
CC	production. The TSARS are easily characterised and have designed
CC	activity allowing direct and rapid detection in a screening process.
CC	Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
QQ	
Query Match	4.2%; Score 38; DB 12; Length 114;
Best Local Similarity	5.4%; Pred. No. 7.23e+09;
Matches	6; Conservative 33; Mismatches 73; Indels 0; Gaps 0;
Dbb	3 bnnbnbnbnbnbtgcnnbn 62
Qy	2490 CTTCCTACGCTTATCGTGCCACAGCGGTGGCATCAGGGGCAAGTCTCTACGTCCAGTG 2549
Dbb	63 bnnbtgcnbnbnbnbnb 114
Qy	2550 CCAGGGGATCCCAGCGGTCTCATCTCTCCAGCGTGTCTTCGACCGCTGTGC 2601
RESULT	6
ID	Q70469 standard; DNA; 114 BP.
DE	Q70469;
DT	07-APR-1995 (first entry)
DN	Generic DNA sequence to generate a random TSAR peptide library.
KW	TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW	effector domain; concatenated heterofunctional protein; linker;
KW	direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
FH	Key
FT	misc_feature
FT	55..60
FT	/*tag= a
FT	/note= "this sequence represents 'Z'; Z can be a

```

sequence of 6,9 or 12 nucleotides (see
comments)"
FT FT WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC (TGC)(NNB)5z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or comps. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 4.1%; Score 37; DB 12; Length 114;
Best Local Similarity 6.3%; Pred. No. 2.78e-08;
Matches 7; Conservative 32; Mismatches 72; Indels 0; Gaps 0;

Db 2 gccnnbnnbnnbnnbnnbnnbnnbgtcgncbnnbnnbnnbnnbnnbnnnnnn 61
Qy 2567 GCTCATCTCTCCACGCTGCTCTGAGCGCTGCTAGCGGCATGTGAGAACAGCTGT 2626

Db 62 nbnnbgtcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbt 112
Qy 2627 TTGCGGGGATTCGGCGGACGGGCTCTCTCTCGTTTGTGGATGATTCT 2677

RESULT 7
ID Q70467 standard; DNA; 114 BP.
DE Q70467;
DT 05-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT FT a
FT FT /note= "this sequence represents 'Z'; Z can be a
FT FT sequence of 6, 9 or 12 nucleotides (see
FT FT comments)"
PD WO9418318-A.
PF 18-AUG-1994.
PR 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI: 94-279739/34
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC (TGC)(NNB)5z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or comps. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

```

```

sequence of 6,9 or 12 nucleotides (see
comments)"
FT
FT WO9418318-A.
PD 18-AUG-1994.
PN
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI WPI: 94-279739/34
DR
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC CC 070469 is a generic DNA sequence used to generate random TSAR peptide
CC (TGC)(NNB)5z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in 070465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or comps. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 4.1%; Score 37; DB 12; Length 114;
Best Local Similarity 6.3%; Pred. No. 2.78e-08;
Matches 7; Conservative 32; Mismatches 72; Indels 0; Gaps 0;

Db 2 gccnnbnnbnnbnnbnnbnnbnnbgtcgcnbnnbnnbnnbnnbnnbnnnnnnnn 61
Qy 2567 GCTCATCTCTCCACGCTGCTCTGAGCGCTGTGCTAGCGGACATGGAGAACAGCTGT 2626
Db 62 nbnnbgtcnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 112
Qy 2627 TTGCGGGGATTCGGCGGACGGGCTCTCTCTCGTTTGTGGATGATTCT 2677

RESULT 7
ID 070467 standard; DNA; 114 BP.
DE 070467;
DT 05-APR-1995 (first entry)
DR Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /tag= a
FT /note= "this sequence represents '2'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PN
PF 01-FEB-1994. U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK;
PI WPI: 94-279739/34
DR
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC CC 070469 is a generic DNA sequence used to generate random TSAR peptide
CC (TGC)(NNB)5z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in 070465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or comps. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

```

```
PR 31-JAN-1994; US-189331.
PA (UYN(-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)16(TGC)(NNB)16(Y). X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 4.0%; Score 36; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.06e-07;
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 2442 CGAGCAGAGCTCCTCCTCAATGAGCGCCAGCAGTGGCCCTCTTCAGCTCTCCTAGCCT 2501
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 2502 CATGTGCCACAGCGCGTGCATCAGGCGCAAGTCTCCTAGCTCCAGTGCCAG 2553
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN(-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
```

```
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 4.0%; Score 36; DB 12; Length 114;
Best Local Similarity 3.6%; Pred. No. 1.06e-07;
Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 2055 CGTGCTCAACTACGAGCGCGCGCGCCCGCTCTCTGCTGCTGCGGCGCT 2114
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 114
QY 2115 GGACGATATCCACAGGCGCGCTTCGTGCTGCTGCGGCGCCAG 2166
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DE Generic DNA sequence to generate a random TSAR-9 peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN(-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows:X(NNB)11(TGC)(NNB)10(TGC)2(NNB)42(NNB)8(TGC)(NNB)
CC -9;. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
```



[illegible]

CC 001, they can also replace the function of macrophages, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed  
CC activity allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP: 5 A; 10 C; 0 G; 0 T;





**THIS PAGE BLANK (USP10)**

\*\*\*\*\*

W P S R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 18:38:12 1998; MasPar time 1550.93 Seconds  
Tabular output not generated. 622.512 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1920-2820) from US08951733.seq (2 of 2)  
Perfect Score: 901  
N.A. Sequence: 1920 GTCAGACTCGGTCATCC.....AGGCCCTGGTGGCAGCGCT 2820  
Comp: CAGGTCGTGAGCGAAGTAGG.....TCCGGGACCCACCGTCCGA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777091 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-pending  
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005  
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82  
15:U83 16:U84A 17:U84B 18:U85 19:U86 20:U87 21:U88  
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8  
29:NEWU9

Statistics: Mean 9.874; Variance 3.555; scale 2.777

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	901	100.0	2848	22	US-08-951-Sequence 13, Applicati	0.00e+00
2	901	100.0	3798	22	US-08-951-Sequence 19, Applicati	0.00e+00
3	901	100.0	4015	24	US-09-052-Sequence 1, Applicati	0.00e+00
4	901	100.0	4015	21	US-08-854-Sequence 224, Applicat	0.00e+00
5	901	100.0	4015	22	US-08-912-Sequence 1, Applicati	0.00e+00
6	901	100.0	4023	23	US-09-026-Sequence 35, Applicati	0.00e+00
7	901	100.0	7029	22	US-08-911-Sequence 1, Applicati	0.00e+00
8	898	99.7	4029	21	US-08-854-Sequence 173, Applicat	0.00e+00
9	898	99.7	4029	21	US-08-851-Sequence 173, Applicat	0.00e+00
10	508	56.4	2176	22	US-08-912-Sequence 3, Applicati	0.00e+00
11	508	56.4	3346	23	US-09-026-Sequence 36, Applicati	0.00e+00
12	508	56.4	3855	22	US-08-911-Sequence 18, Applicati	0.00e+00
13	508	56.4	3855	22	US-08-912-Sequence 4, Applicati	0.00e+00
14	477	52.9	2171	21	US-08-851-Sequence 100, Applicat	0.00e+00
15	477	52.9	2171	21	US-08-854-Sequence 100, Applicat	0.00e+00

16	477	52.9	2171	21	US-08-846-Sequence 100, Applicat	0.00e+00
17	369	41.0	3496	23	US-09-042-Sequence 1, Applicati	0.00e+00
18	277	30.7	535	22	US-08-911-Sequence 52, Applicati	3.09e-279
19	236	26.2	550	22	US-08-911-Sequence 54, Applicati	1.27e-231
20	182	20.2	182	22	US-08-912-Sequence 9, Applicatio	1.82e-169
21	159	17.6	389	21	US-08-846-Sequence 62, Applicati	2.84e-143
22	159	17.6	389	21	US-08-846-Sequence 62, Applicati	2.84e-143
23	159	17.6	389	21	US-08-851-Sequence 62, Applicati	2.84e-143
24	159	17.6	389	22	US-08-911-Sequence 17, Applicati	2.84e-143
25	159	17.6	389	22	US-08-854-Sequence 62, Applicati	2.84e-143
26	159	17.6	389	22	US-08-912-Sequence 8, Applicatio	2.84e-143
27	66	7.3	578	19	US-08-624-Sequence 14, Applicati	3.75e-41
28	66	7.3	578	18	US-08-569-Sequence 14, Applicati	3.75e-41
29	57	6.3	383	16	US-08-446-Sequence 8, Applicatio	5.74e-32
30	57	6.3	383	16	US-08-446-Sequence 8, Applicatio	5.74e-32
31	57	6.3	383	16	US-08-446-Sequence 8, Applicatio	5.74e-32
32	57	6.3	383	16	US-08-446-Sequence 8, Applicatio	5.74e-32
33	57	6.3	383	17	US-08-451-Sequence 8, Applicatio	5.74e-32
34	57	6.3	383	15	US-08-311-Sequence 8, Applicatio	5.74e-32
35	55	6.1	383	16	US-08-446-Sequence 8, Applicatio	5.81e-30
36	55	6.1	383	16	US-08-446-Sequence 8, Applicatio	5.81e-30
37	55	6.1	383	16	US-08-446-Sequence 8, Applicatio	5.81e-30
38	55	6.1	383	17	US-08-451-Sequence 8, Applicatio	5.81e-30
39	55	6.1	383	16	US-08-446-Sequence 8, Applicatio	5.81e-30
40	55	6.1	383	15	US-08-311-Sequence 8, Applicatio	5.81e-30
41	54	6.0	60	22	US-08-911-Sequence 44, Applicati	5.77e-29
42	54	6.0	60	22	US-08-912-Sequence 328, Applicati	5.77e-29
43	54	6.0	477	11	US-07-904-Sequence 10, Applicati	5.77e-29
44	53	5.9	239	22	US-08-911-Sequence 53, Applicati	5.68e-28
45	53	5.9	578	19	US-08-624-Sequence 14, Applicati	5.68e-28

ALIGNMENTS

RESULT 1  
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.  
AC xxxxxx  
DT

Sequence 13, Application US/08951733  
Sequence 13, Application US/08951733  
GENERAL INFORMATION:

APPLICANT: Harrington, Lea A.

APPLICANT: Robinson, Murray O.

TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: One Angen Center Drive

CITY: Thousand Oaks

STATE: CA

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/951,733

FILING DATE: 16-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/873,039

FILING DATE: 11-JUN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/751,189

FILING DATE: 15-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Oleski, Nancy A.

REGISTRATION NUMBER: 34,688

REFERENCE/DOCKET NUMBER: A-433B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2848 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
SQ SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 22; Length 2848;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1979  
Qy 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1979

Db 1980 CFTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039  
Qy 1980 CFTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039

Db 2040 GAAGGCACTGTTACGGTGTCTCAACTACAGCGGGCGGCCGCCGCTCTCTGGGCGC 2099  
Qy 2040 GAAGGCACTGTTACGGTGTCTCAACTACAGCGGGCGGCCGCCGCTCTCTGGGCGC 2099

Db 2100 CFTCTGTGTCGGCCCTGGAGATATCCACAGGGCCTGGCGACCTTCGTGCTGTGGTGG 2159  
Qy 2100 CFTCTGTGTCGGCCCTGGAGATATCCACAGGGCCTGGCGACCTTCGTGCTGTGGTGG 2159

Db 2160 GCGCCAGGACCGCGCCCTGAGCTGTACTTTGTCAGGTGGATGTGAGGGGGCGGTACGA 2219  
Qy 2160 GCGCCAGGACCGCGCCCTGAGCTGTACTTTGTCAGGTGGATGTGAGGGGGCGGTACGA 2219

Db 2220 CACCATCCCCAGGACAGCTCACGGAGTTCATCCGACGATCATCAACCCAGAACAC 2279  
Qy 2220 CACCATCCCCAGGACAGCTCACGGAGTTCATCCGACGATCATCAACCCAGAACAC 2279

Db 2280 GTACTGCGTGTGCGTATGCGGTGCGTCCAGAAAGGCGGCCATGGGACGTCGCAAGGC 2339  
Qy 2280 GTACTGCGTGTGCGTATGCGGTGCGTCCAGAAAGGCGGCCATGGGACGTCGCAAGGC 2339

Db 2340 CTTCAAGAGCCACGCTCTACCTTCACAGACCTCCAGCGGTACATGCCACAGTTGCTGGC 2399  
Qy 2340 CTTCAAGAGCCACGCTCTACCTTCACAGACCTCCAGCGGTACATGCCACAGTTGCTGGC 2399

Db 2400 TCACCTGCAGGACAGACCGCCGCTGAGGGATGCCGTGCTATCGAGCAGAGCTCTCCCT 2459  
Qy 2400 TCACCTGCAGGACAGACCGCCGCTGAGGGATGCCGTGCTATCGAGCAGAGCTCTCCCT 2459

Db 2460 GAATGAGGCCAGAGTGGCCCTTTCCAGCTTTCTTACGCTTCATGTCGACGAGCGCGT 2519  
Qy 2460 GAATGAGGCCAGAGTGGCCCTTTCCAGCTTTCTTACGCTTCATGTCGACGAGCGCGT 2519

Db 2520 GCGCATCAGGGCCAGTCTTAGTCCAGTCCAGGGATCCCGAGGGCTCCATCCTCTC 2579  
Qy 2520 GCGCATCAGGGCCAGTCTTAGTCCAGTCCAGGGATCCCGAGGGCTCCATCCTCTC 2579

Db 2580 CACGCTGCTCTCAGCCTGTGCTACGGGCACATGGAGAACAGCTGTTTGGGGGATTGG 2639  
Qy 2580 CACGCTGCTCTCAGCCTGTGCTACGGGCACATGGAGAACAGCTGTTTGGGGGATTGG 2639

Db 2640 GCGGACCGGGCTGCTCCTCGGTTTGGTGGATGATTTCTTGTGTCACACCTCACCTCAC 2699  
Qy 2640 GCGGACCGGGCTGCTCCTCGGTTTGGTGGATGATTTCTTGTGTCACACCTCACCTCAC 2699

Db 2700 CCACCGGAAACCTTCTCAGACCCCTGGTCGAGGTGTCCTGAGTATGGCTGGCTGGT 2759  
Qy 2700 CCACCGGAAACCTTCTCAGACCCCTGGTCGAGGTGTCCTGAGTATGGCTGGCTGGT 2759

Db 2760 GAACTTGGGAAACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGC 2819  
Qy 2760 GAACTTGGGAAACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGC 2819

Db 2820 T 2820  
Qy 2820 T 2820

RESULT 2  
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.  
AC xxxxxx  
DT  
DE Sequence 19, Application US/08951733  
CC Sequence 19, Application US/08951733  
CC GENERAL INFORMATION:  
CC APPLICANT: Harrington, Lea A.  
CC APPLICANT: Robinson, Murray O.  
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
CC NUMBER OF SEQUENCES: 44  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Amgen Inc.  
CC STREET: One Amgen Center Drive  
CC CITY: Thousand Oaks  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 91320-1789  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/951,733  
CC FILING DATE: 16-OCT-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/873,039  
CC FILING DATE: 11-JUN-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/751,189  
CC FILING DATE: 15-NOV-1996  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Oleski, Nancy A.  
CC REGISTRATION NUMBER: 34,688  
CC REFERENCE/DOCKET NUMBER: A-433B  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (805) 447-6504  
CC TELEFAX: (805) 499-8011  
CC INFORMATION FOR SEQ ID NO: 19:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3798 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
SQ SEQUENCE 3798 BP; 613 A; 1310 C; 1213 G; 562 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 22; Length 3798;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1971 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1980  
Qy 1971 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGTGCGGCCGATTGTGAACATGGACTA 1979

Db 1981 CGTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2040  
Qy 1981 CGTCTGGGAGCCAGACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGTT 2039

Db 2041 GAAGGCACTGTTACGCTGTCTCAACTACGAGGGCGGCCGCCGCTCTCTGGGCGC 2100  
Qy 2041 GAAGGCACTGTTACGCTGTCTCAACTACGAGGGCGGCCGCCGCTCTCTGGGCGC 2099

Db 2101 CTCTGTGCTGGGCTTGACGATATCCACAGGGCTGCGGACCTTCGTGCTGCGTGGCG 2160  
Qy 2101 CTCTGTGCTGGGCTTGACGATATCCACAGGGCTGCGGACCTTCGTGCTGCGTGGCG 2160

QY 2100 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGGCGCACCTTCGTGCTGCTGTGG 2159  
Db 2161 GGCCAGGACCCCGCTGAGCTGACTTTGTTCAGAGTGGATGTGACGGCGCGTACGA 2220  
QY 2160 GGCCAGGACCCCGCTGAGCTGACTTTGTTCAGAGTGGATGTGACGGCGCGTACGA 2219  
Db 2221 CACATCCCCCAGGACAGGCTCAGGAGGTGTCATGCCAGCATCATCAAAACCCAGAACAC 2280  
QY 2220 CACATCCCCCAGGACAGGCTCAGGAGGTGTCATGCCAGCATCATCAAAACCCAGAACAC 2279  
Db 2281 GTACTGCGTGGCTGCTATGCCGTGTGCTCCAGAGGGCGCCATGGGCACCTCCGCAAGGC 2340  
QY 2280 GTACTGCGTGGCTGCTATGCCGTGTGCTCCAGAGGGCGCCATGGGCACCTCCGCAAGGC 2339  
Db 2341 CTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAGCGGTATAGTGGAGTTCGTGGC 2400  
QY 2340 CTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAGCGGTATAGTGGAGTTCGTGGC 2399  
Db 2401 TCACCTGCAGGAGACAGCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCT 2460  
QY 2400 TCACCTGCAGGAGACAGCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCT 2459  
Db 2461 GAATGAGGCCAGCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTGCCACCAACGCGT 2520  
QY 2460 GAATGAGGCCAGCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTGCCACCAACGCGT 2519  
Db 2521 GCGCATCAGGGGAAGTCTTACGTTCAGTCCAGGGGATCCCGAGGGCTCCATCTCTC 2580  
QY 2520 GCGCATCAGGGGAAGTCTTACGTTCAGTCCAGGGGATCCCGAGGGCTCCATCTCTC 2579  
Db 2581 CAGCTGCTCTGAGCGCTGTGTCAGCGGACATGGAGAACAGCTGTTCCGGGGATTGG 2640  
QY 2580 CAGCTGCTCTGAGCGCTGTGTCAGCGGACATGGAGAACAGCTGTTCCGGGGATTGG 2639  
Db 2641 GCGGGAGCGGCTGCTGCGGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCAC 2700  
QY 2640 GCGGGAGCGGCTGCTGCGGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCAC 2699  
Db 2701 CCACGGGAAAACCTTCTCAGGACCTGTCGAGGTGTCGAGTATGCGTGGT 2760  
QY 2700 CCACGGGAAAACCTTCTCAGGACCTGTCGAGGTGTCGAGTATGCGTGGT 2759  
Db 2761 GAACCTGCGGAAGACAGTGTGTAACCTCCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820  
QY 2760 GAACCTGCGGAAGACAGTGTGTAACCTCCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2819  
Db 2821 T 2821  
QY 2820 T 2820

RESULT 3  
ID US-09-052-919-1 STANDARD; DNA; UNC; 4015 BP.  
AC xxxxxx  
DT

Sequence 1, Application US/09052919  
Sequence 1, Application US/09052919  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Antisense Compositions for Detecting and  
TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California

CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/052,919  
CC FILING DATE: 31-MAR-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,549  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,584  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Parent, Annette S.  
CC REGISTRATION NUMBER: 42,058  
CC REFERENCE/DOCKET NUMBER: 015389-00360005  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4015 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56..3454  
CC OTHER INFORMATION: /product= "human telomerase reverse  
CC OTHER INFORMATION: transcriptase (hTERT)"  
CC SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

Query Match 100.0%; Score 901; DB 24; Length 4015;  
Best Local Similarity 100.0%; Pred.No. 0.00e+00;  
Matches 901; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1909 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTGCGCCGATTGTGAACATGGACTA 1968









Db	2332	CTTCAAGAGCCACGTCCTACCTTGACAGACCTCAGCCGTACATGGACAGATTCGTGGC	2391
QY	2340	CTTCAAGAGCCACGTCCTACCTTGACAGACCTCAGCCGTACATGGACAGATTCGTGGC	2399
Db	2392	TCACCTGTCAGGAGACCAAGCCGCTGAGGGATGCGCTCGTCATCGAGCAGAGCTCCTCCCT	2451
QY	2400	TCACCTGTCAGGAGACCAAGCCGCTGAGGGATGCGCTCGTCATCGAGCAGAGCTCCTCCCT	2459
Db	2452	GAATGAGGCCAGAGTGGCCTCTTCGACGCTCTTCTACGCTTCAATGCTGCCACCAACGCCGT	2511
QY	2460	GAATGAGGCCAGAGTGGCCTCTTCGACGCTCTTCTACGCTTCAATGCTGCCACCAACGCCGT	2519
Db	2512	GCGCATCAGGGCAAGTCTACGTCAGTGCAGGGGATCCGCAGAGGCTCCATCCTCTC	2571
QY	2520	GCGCATCAGGGCAAGTCTACGTCAGTGCAGGGGATCCGCAGAGGCTCCATCCTCTC	2579
Db	2572	CACGCTGCTCTGCAGCCTGTGCTACGGGGACATGGAGAAACAGCTGTTGCGGGGATTTCG	2631
QY	2580	CACGCTGCTCTGCAGCCTGTGCTACGGGGACATGGAGAAACAGCTGTTGCGGGGATTTCG	2639
Db	2632	CGGGAGCGGCTGCTCCTCGGTTTGGTGGATGATTTCTTGTGTGACAACTCACCTCAC	2691
QY	2640	CGGGAGCGGCTGCTCCTCGGTTTGGTGGATGATTTCTTGTGTGACAACTCACCTCAC	2699
Db	2692	CCACGCGAAAACCTTCCTCAGACACCTGGTCCGAGGTCTCCCTGAGTATGGCTCCGTGGT	2751
QY	2700	CCACGCGAAAACCTTCCTCAGACACCTGGTCCGAGGTCTCCCTGAGTATGGCTCCGTGGT	2759
Db	2752	GAACTTCGCGAAGACAGTGGTGAATTCCTCTGTAGAAGACGAGGCCCTGGTGGCAGCGC	2811
QY	2760	GAACTTCGCGAAGACAGTGGTGAATTCCTCTGTAGAAGACGAGGCCCTGGTGGCAGCGC	2819
Db	2812	T 2812	
QY	2820	T 2820	

RESULT 7  
ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.  
AC xxxxxx

DE	Sequence 1, Application US/08911312
CC	Sequence 1, Application US/08911312

GENERAL INFORMATION:

CC GENERAL INFORMATION:  
CC APPLICANT: Cecb. Thomas R

CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Finney, Joachim

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC  
APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC  
APPLICANT: Andrews, William

CC AFFILIAN. ANDREWS, WILLIAM  
CC TITLE OF INVENTION: TEL

CC	TITLE OF INVENTION:	TELEPHONE
CC	NUMBER OF SEQUENCES:	170

CC NUMBER OF SEQUENCES: 170

CC CORRESPONDENCE ADDRESS:

CC ADDRESS: Townsend and

CC STREET: Two Embarcadero

CC CITY: San Francisco

STATE: California

STATE: CALIFORNIA  
COUNTRY: USA

COUNTRY: USA  
ZIP: 04111-2834

CC  
CC  
ZIP: 94111-3834

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy di:

CC  
COMPUTER: IBM PC compa

**COMPUTER:** IBM PC COMP  
**OPERATING SYSTEM:** PC-DOS

CC OF EXAMINING SYSTEM: Patent In Rel  
CC SOFTWARE: Patent In Rel

```

CC      SOFTWARE:  PATENTLTH REL
CC
CC      CURRENT APPLICATION DATA:
CC

```

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US

CC FILING DATE: 14-AUG-1991

Db	2332	CTTCAAGAGCCACGTCCTTACCTTCACAGACCTCCAGCCGTACATCGCACA
QY	2340	CTTCAAGAGCCACGTCCTTACCTTCACAGACCTCCAGCCGTACATCGCACA
Db	2392	TCACCTGCAGGAGACACGACCCGCTCAGGGATGCGGTCTCATCTGACGACAGAG
QY	2400	TCACCTGCAGGAGACACGACCCGCTCAGGGATGCGGTCTCATCTGACGACAGAG
Db	2452	GAATGAGGCCAGCACTGCGCTCTTCGACGCTTCTTCTACGCTTTCATGTGCCACA
QY	2460	GAATGAGGCCAGCACTGCGCTCTTCGACGCTTCTTCTACGCTTTCATGTGCCACA
Db	2512	CGCATCAGGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGCAGGGGCTCT
QY	2520	CGCATCAGGGGCAAGTCTCTACGTCCAGTCCAGGGGATCCCGCAGGGGCTCT
Db	2572	CACGCTGCTCTGCAGCCCTGTGCTACGGCGACATGGAGAACAAGCTCTTTGCC
QY	2580	CACGCTGCTCTGCAGCCCTGTGCTACGGCGACATGGAGAACAAGCTCTTTGCC
Db	2632	CGGGACGGGCTGCTCTCGCTTTGGTGGATGATTTCTTTGGTGCACACCTTGG
QY	2640	CGGGACGGGCTGCTCTCGCTTTGGTGGATGATTTCTTTGGTGCACACCTTGG
Db	2692	CCAGCGAAAAACCTTCTCTCAGACCCCTGGTCCGAGGTGCCCTGAGTATGAGG
QY	2700	CCAGCGAAAAACCTTCTCTCAGACCCCTGGTCCGAGGTGCCCTGAGTATGAGG
Db	2752	GAATTCGGGAGACAGTGGTCAACTTCCCTGTAGNAGACGAGGCCCTGGG
QY	2760	GAATTCGGGAGACAGTGGTCAACTTCCCTGTAGNAGACGAGGCCCTGGG
Db	2812	T 2812
QY	2820	T 2820
RESULT	7	
ID	US-08-911-312-1	STANDARD; DNA; UNC; 7029 BP.
AC	xxxxxx	
DT		
DE	Sequence 1, Application US/08911312	
CC	Sequence 1, Application US/08911312	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Cech, Thomas R.	
CC	APPLICANT: Lingner, Joachim	
CC	APPLICANT: Nakamura, Toru	
CC	APPLICANT: Chapman, Karen B.	
CC	APPLICANT: Morin, Gregg B.	
CC	APPLICANT: Harley, Calvin B.	
CC	APPLICANT: Andrews, William	
CC	TITLE OF INVENTION: Telomerase Reverse Transcriptase	
CC	NUMBER OF SEQUENCES: 170	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Townsend and Townsend and Crew LLP	
CC	STREET: Two Embarcadero Center, Eighth Floor	
CC	CITY: San Francisco	
CC	STATE: California	
CC	COUNTRY: USA	
CC	ZIP: 94111-3834	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/08/911,312	
CC	FILING DATE: 14-AUG-1997	
CC	CLASSIFICATION: 536	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 08/724,643	
CC	FILING DATE: 01-OCT-1996	

Db 3055 CTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGGC 3114  
QY 2340 CTTCAAGAGCCAGCTCTTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGGC 2399  
Db 3115 TCACCTGCAGGAGACAGCCCGCTGAGGATGCGCTGCTCATCGAGCAGAGCTCTCCCT 3174  
QY 2400 TCACCTGCAGGAGACAGCCCGCTGAGGATGCGCTGCTCATCGAGCAGAGCTCTCCCT 2459  
Db 3175 GAATGAGCCAGCAGTGGCTTTCGAGCTTTCCTAGCTTTCCTAGCTTTCATGTGCCACAGCGCT 3234  
QY 2460 GAATGAGCCAGCAGTGGCTTTCGAGCTTTCCTAGCTTTCCTAGCTTTCATGTGCCACAGCGCT 2519  
Db 3235 GCGCATCAGGGCAGTCTTCTAGTCCATGCGCAGGGATCCCGAGGCTCCATCTCTC 3294  
QY 2520 GCGCATCAGGGCAGTCTTCTAGTCCATGCGCAGGGATCCCGAGGCTCCATCTCTC 2579  
Db 3295 CACGCTGCTCTGACGCTTCTGTCAGGCGACATGAGAACAGCTGTTTGGGGGATTCG 3354  
QY 2580 CACGCTGCTCTGACGCTTCTGTCAGGCGACATGAGAACAGCTGTTTGGGGGATTCG 2639  
Db 3355 GCGGACCGGCTGCTCTGCTGCTTGGTGATGATTTCTTGTGTGACACCTCACCTCAC 3414  
QY 2640 GCGGACCGGCTGCTCTGCTGCTTGGTGATGATTTCTTGTGTGACACCTCACCTCAC 2699  
Db 3415 CCACCGAAACCTTCTCCTCAGACCTCGTCCGAGGTCTCCCTGATGATGCTGGTGGT 3474  
QY 2700 CCACCGAAACCTTCTCCTCAGACCTCGTCCGAGGTCTCCCTGATGATGCTGGTGGT 2759  
Db 3475 GAACCTGGGAGAGAGTGGTGAATCTCCCTGTAGAAGACGAGGCGCTGGTGACACGGC 3534  
QY 2760 GAACCTGGGAGAGAGTGGTGAATCTCCCTGTAGAAGACGAGGCGCTGGTGACACGGC 2819  
Db 3535 T 3535  
QY 2820 T 2820

RESULT 8

ID US-08-854-050-173 STANDARD; DNA; UNC; 4029 BP.  
AC xxxxxx

Sequence 173, Application US/08854050  
Sequence 173, Application US/08854050

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Novel Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-00293005  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 173:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4029 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: CDNA  
CC FEATURE:  
CC NAME/KEY:  
CC LOCATION: 1..4029  
CC OTHER INFORMATION: /note= "preliminary sequence for  
CC OTHER INFORMATION: human TRT CDNA insert of  
CC OTHER INFORMATION: plasmid pGRI121"  
CC SQ SEQUENCE 4029 BP; 687 A; 1342 C; 1255 G; 720 T; 25 OTHER.

Query Match 99.7%; Score 898; DB 21: Length 4029;  
Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
Matches 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1904 GTCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1963  
QY 1920 GTCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1979  
Db 1964 CGTCGTGGGAGCCAGAACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTACCTCGAGGGT 2023  
QY 1980 CGTCGTGGGAGCCAGAACGTTCCGACAGAGAAAGAGGGCCGAGCGTCTACCTCGAGGGT 2039  
Db 2024 GAAGGCACCTGTTGAGCGTCTCAACTACGAGCGGGCGGGCCCGCCCTCTCTGGGCGC 2083  
QY 2040 GAAGGCACCTGTTGAGCGTCTCAACTACGAGCGGGCGGGCCCGCCCTCTCTGGGCGC 2099  
Db 2084 CTCTGTCTGGGCTGGACGATATCCACAGGGCTCGGGCCACCTTCTGCTGGGTGTCG 2143  
QY 2100 CTCTGTCTGGGCTGGACGATATCCACAGGGCTCGGGCCACCTTCTGCTGGGTGTCG 2159  
Db 2144 GGCCACAGACCCGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2203  
QY 2160 GGCCACAGACCCGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2219  
Db 2204 CACATCCCCCAGGACAGGCTCAGGGAGTTCATCGCCAGCATCATCAAAACCCAGAACAC 2263  
QY 2220 CACATCCCCCAGGACAGGCTCAGGGAGTTCATCGCCAGCATCATCAAAACCCAGAACAC 2279  
Db 2264 GTACTGCGTGGTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2323  
QY 2280 GTACTGCGTGGTGGTATGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339  
Db 2324 CTTCAAGAGCCAGCTCTCTACCTTTGACAGACCTCCAGCCCTACATGCGACAGTTGCTGGC 2383  
QY 2340 CTTCAAGAGCCAGCTCTCTACCTTTGACAGACCTCCAGCCCTACATGCGACAGTTGCTGGC 2399  
Db 2384 TCACCTGTCAGGAGAACAGCCCGCTGAGGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2443  
QY 2400 TCACCTGTCAGGAGAACAGCCCGCTGAGGGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2459



QY 2580 CACGCTGCTGCAGCCCTGCTAGCGGACATGGAGACAAAGCTGTTTGGGGGATTGC 2639

Db 2624 GCGGACGGGCTGCTCCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2683

QY 2640 GCGGACGGGCTGCTCCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2699

Db 2684 CACGCGAAACCTTCTCAGACCTCGTCCGAGGTGCTCCTGAGTATGCTGGTGGT 2743

QY 2700 CACGCGAAACCTTCTCAGACCTCGTCCGAGGTGCTCCTGAGTATGCTGGTGGT 2759

Db 2744 GAACCTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGTGCGACGGC 2803

QY 2760 GAACCTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGTGCGACGGC 2819

Db 2804 T 2804

QY 2820 T 2820

RESULT 10

AC US-08-912-951-3 STANDARD; DNA; UNC; 2176 BP.

DT xxxxxx

Sequence 3, Application US/08912951

Sequence 3, Application US/08912951

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/912,951

FILING DATE: 14-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

CC NAME: Apple, Randolph T.

CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002600US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200

CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 3:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 2176 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cdna

CC FEATURE:

CC NAME/KEY: -

CC LOCATION: 1..2176

CC OTHER INFORMATION: /note= "clone 712562"

CC SEQUENCE 2176 BP; 432 A; 678 C; 642 G; 422 T; 2 OTHER.

Query Match 56.4%; Score 508; DB 22; Length 2176;

Best Local Similarity 79.8%; Pred. No. 0.00e+00;

Matches 719; Conservative 0; Mismatches 0; Indels 182; Gaps 1;

Db 232 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTTGTGAACATGGACTA 291

QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTTGTGAACATGGACTA 1979

Db 292 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGGCCGAGCGTCTCACCTCGAGGGT 351

QY 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAGAGGCCGAGCGTCTCACCTCGAGGGT 2039

Db 352 GAAGGCACCTGTTCAAGCGTGCTCAACTACGAGCGGGCGCGCCCGCCCTCTCTGGCGC 411

QY 2040 GAAGGCACCTGTTCAAGCGTGCTCAACTACGAGCGGGCGCGCCCGCCCTCTCTGGCGC 2099

Db 412 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGCGGCACCTTCTGCTGGGTGCG 471

QY 2100 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGCGGCACCTTCTGCTGGGTGCG 2159

Db 472 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGA 531

QY 2160 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGA 2219

Db 532 CACATCCCCAGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAAAC 591

QY 2220 CACATCCCCAGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCAGAAAC 2279

Db 592 GTACTGCGTGGTGGTATGCGGTGTCAGAGAGGCCGCCCATGGGCACGTCCGCAAGGC 651

QY 2280 GTACTGCGTGGTGGTATGCGGTGTCAGAGAGGCCGCCCATGGGCACGTCCGCAAGGC 2339

Db 652 CTTCAAGAGCCACGTC----- 667

QY 2340 CTTCAAGAGCCACGTCCTTACCTTGACAGACCTCCAGCCCGTACATGCGACAGTTCGTGGC 2399

Db 667 ----- 667

QY 2400 TCACCTGCAGGAGACACGCGCGCTGAGGGATGCGGTGCTCATGCGAGCAGAGCTCCTCCCT 2459

Db 667 ----- 667

QY 2460 GAATGAGCCAGCAGTGGCCTTTCGACGTCCTTCCAGCTTCATGTGCCACACGCGCT 2519

Db 668 -----CTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCCTCTC 709

QY 2520 GCGCATCAGGGGCAAGTCTTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCCTCTC 2579

Db 710 CACGCTGCTGACGCGCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 769

QY 2580 CACGCTGCTGACGCGCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 2639

Db 770 GCGGAGCGGCTGCTCCTGCTGGTGGATGATTTCTTGTGTGACACTCACCTCAC 829

QY 2640 GCGGACGGGCTGCTCTCGCTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2699  
Db 830 CCACGCGAAACCTTCCTCAGACACCTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGGT 889  
QY 2700 CCACGCGAAACCTTCCTCAGACACCTGGTCCGAGGTGTCCCTGAGTATGGCTGGTGGT 2759  
Db 890 GAACCTGGGGAAGACAGTGGTGAACCTCCCTGTAGAACAGAGGCGCTGGTGCGACGCG 949  
QY 2760 GAACCTGGGGAAGACAGTGGTGAACCTCCCTGTAGAACAGAGGCGCTGGTGCGACGCG 2819  
Db 950 T 950  
QY 2820 T 2820

## RESULT 11

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.  
AC xxxxxx  
DT  
DE Sequence 36, Application US/09026981  
CC Sequence 36, Application US/09026981  
CC GENERAL INFORMATION:  
CC APPLICANT: Counter, Christopher M.  
CC APPLICANT: Meyerson, Matthew  
CC APPLICANT: Weinberg, Robert A.  
CC TITLE OF INVENTION: Telomerase Catalytic Subunit Gene and  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
CC STREET: Two Militia Drive  
CC CITY: Lexington  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02173  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/026.981  
CC FILING DATE: 20-FEB-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/064,322  
CC FILING DATE: 30-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/055,762  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/054,549  
CC FILING DATE: 01-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/047,151  
CC FILING DATE: 20-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 60/038,750  
CC FILING DATE: 20-FEB-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Granahan, Patricia  
CC REGISTRATION NUMBER: 32,227  
CC REFERENCE/DOCKET NUMBER: WHI97-11p4AM  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 781-861-6240  
CC TELEFAX: 781-861-9540  
CC INFORMATION FOR SEQ ID NO: 36:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3346 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC SEQUENCE 3346 BP; 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 56.4%; Score 508; DB 23; Length 3346;  
Best Local Similarity 79.8%; Pred. No. 0.00e+00;  
Matches 719; Conservative 0; Mismatches 0; Indels 182; Gaps 1;  
Db 1251 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1310  
QY 1920 GTCCAGACTCCGCTTCATCCCAAGCCTGACGGGCTCGGGCCGATTGTGAACATGGACTA 1979  
Db 1311 CGTCGTGGGAGCCAGAACGTTCCGACAGAAAAGAGGGCCGAGCGTCTCACTCGAGGGT 1370  
QY 1980 CGTCGTGGGAGCCAGAACGTTCCGACAGAAAAGAGGGCCGAGCGTCTCACTCGAGGGT 2039  
Db 1371 GAAGGCACCTGTTACGCGTCTCAACTACGACGGGCGGGGCGGCCCTCCCTGGGGCG 1430  
QY 2040 GAAGGCACCTGTTACGCGTCTCAACTACGACGGGCGGGGCGGCCCTCCCTGGGGCG 2099  
Db 1431 CTCTGTGCTGGGCGCTGACGATATCCACAGGGCCTGGCGACCTTCGTGTGCTGTGCG 1490  
QY 2100 CTCTGTGCTGGGCGCTGACGATATCCACAGGGCCTGGCGACCTTCGTGTGCTGTGCG 2159  
Db 1491 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGCTACGA 1550  
QY 2160 GGCCAGGACCCGCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGCTACGA 2219  
Db 1551 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCCAAGACAC 1610  
QY 2220 CACCATCCCCAGGACAGGCTCAGGAGGTATCGCCAGCATCATCAAAACCCCAAGACAC 2279  
Db 1611 GTACTGCTGCTGGTATGCGGTGTCGAAAGGCGGCCCATGGCAGCGTCCGCAAGGC 1670  
QY 2280 GTACTGCTGCTGGTATGCGGTGTCGAAAGGCGGCCCATGGCAGCGTCCGCAAGGC 2339  
Db 1671 CTTCAAGAGCCACGTC----- 1686  
QY 2340 CTTCAAGAGCCACGTCCTACTTTGACAGACCTCCAGCGGTACATGCGACAGTTGCGTGGC 2399  
Db 1686 ----- 1686  
QY 2400 TCACCTGCAGAGACACGCGCGTGGAGGATGCCGTGCTCATCGAGCAGAGCTCCTCCCT 2459  
Db 1686 ----- 1686  
QY 2460 GAATGAGCCAGCAGTGGCGCTCTTCGAGCTTCTCCTAGCTTCATGTGCCACACGCGGT 2519  
Db 1687 -----CTACGTCCAGTGCAGGGATCCCGCAGGGCTCCATCCTCTC 1728  
QY 2520 GCGCATCAGGGGCAAGTCCCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCCTCTC 2579  
Db 1729 CACGCTGCTGCAAGCCTGTGCTACGGGACATGGAGAACAGCTGTTGGGGGATTCG 1788  
QY 2580 CACGCTGCTGCAAGCCTGTGCTACGGGACATGGAGAACAGCTGTTGGGGGATTCG 2639  
Db 1789 GCGGAGCGGCTGCTGCTGGGTTGGTGGATGATTTCTTGTGTGACACCTCACCTCAC 1848  
QY 2640 GCGGAGCGGCTGCTGCTGGGTTGGTGGATGATTTCTTGTGTGACACCTCACCTCAC 2699  
Db 1849 CCACGCGAAACCTTCCTCAGGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGT 1908  
QY 2700 CCACGCGAAACCTTCCTCAGGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGT 2759  
Db 1909 GAACCTGGGGAAGACAGTGGTGAACCTTCCTGTAGAACAGAGGCGCTGGGTGGCAGGCG 1968  
QY 2760 GAACCTGGGGAAGACAGTGGTGAACCTTCCTGTAGAACAGAGGCGCTGGGTGGCAGGCG 2819  
Db 1969 T 1969  
QY 2820 T 2820

## RESULT 12

ID US-08-911-312-18 STANDARD; DNA; UNC; 3855 BP.  
AC xxxxxx  
DT







APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
THERAPEUTIC METHODS  
NUMBER OF SEQUENCES: 335  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3855 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3855 /note= "nucleic acid sequence with an  
OTHER INFORMATION: open reading frame encoding a delta-182  
OTHER INFORMATION: variant polypeptide"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..2479 /product= "delta-182 variant  
OTHER INFORMATION: polypeptide"  
SEQUENCE 3855 BP; 651 A; 1300 C; 1226 G; 678 T; 0 OTHER.

Query Match 56.4%; Score 508; DB 22; Length 3855;  
Best Local Similarity 79.8%; Pred. No. 0.00e+00;  
Matches 719; Conservative 0; Mismatches 0; Indels 182; Gaps 1;  
Db 1909 GTCCAGACTCGCGCTTATCCCAAGCCCTGAGGGGCTGGGGCGGATTTGTGAACATGGACTA 1968  
|||||  
Qy 1920 GTCCAGACTCGCGCTTATCCCAAGCCCTGAGGGGCTGGGGCGGATTTGTGAACATGGACTA 1979  
Db 1969 CGTCTGGAGCCAGACGTTTCCCGAGAGAAAGAGGGCCGAGCGTCTCAGCTCAGAGGT 2028  
|||||  
Qy 1980 CGTCTGGAGCCAGACGTTTCCCGAGAGAAAGAGGGCCGAGCGTCTCAGCTCAGAGGT 2039  
Db 2029 GAAGGCACTGTTACAGCTGCTCAACTAGAGGGGCGGGCGGGCCCGCCCTCTGGGGCGC 2088  
|||||  
Qy 2040 GAAGGCACTGTTACAGCTGCTCAACTAGAGGGGCGGGCGGGCCCGCCCTCTGGGGCGC 2099  
Db 2089 CTCTGTCTGGGCTTGAGACGATATCCACAGGGCCCTGGCGCACCTTCGTGCTGCGTGGG 2148  
|||||  
Qy 2100 CTCTGTCTGGGCTTGAGACGATATCCACAGGGCCCTGGCGCACCTTCGTGCTGCGTGGG 2159  
Db 2149 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGTGGATGTGACGGGCGCGTACGA 2208  
|||||  
Qy 2160 GGCCAGGACCGCGCGCTGAGCTGTACTTTGTCAAGTGGATGTGACGGGCGCGTACGA 2219  
Db 2209 CACCATCCCGCAGGACAGGCTCACGGAGGTCTATCCAGAGCATCAAAACCCCAAGAAC 2268  
|||||  
Qy 2220 CACCATCCCGCAGGACAGGCTCACGGAGGTCTATCCAGAGCATCAAAACCCCAAGAAC 2279  
Db 2269 GTACTGCTGCGTGGTATGCCGTGTCAGAAAGCCGCCCATGGGACGCTCCGCAAGGC 2328  
|||||  
Qy 2280 GTACTGCTGCGTGGTATGCCGTGTCAGAAAGCCGCCCATGGGACGCTCCGCAAGGC 2339  
Db 2329 CTTCAAGAGCCACGTC----- 2344  
|||||  
Qy 2340 CTTCAAGAGCCACGTCCTCTACCTTCACAGACCTCCAGCGGTACATCGACAGTTCGTGGC 2399  
Db 2344 ----- 2344  
Qy 2400 TCACCTGCAGGAGACCAAGCCCGCTGAGGGATGCCGTGTCATCGAGAGAGTCTCCTCC 2459  
Db 2344 ----- 2344  
Qy 2460 GAATGAGCCAGCAGTGGCTCTTCGACGCTTCCTAGCTTCATGTGCCACACGCCGT 2519  
Db 2345 -----CTAGCTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTC 2386  
|||||  
Qy 2520 GGCATCAGGGGCAAGTCTTACGTCCAGTGCAGGGGATCCGACAGGGCTCCATCCTCTC 2579  
Db 2387 CACGCTGCTGACGCCCTGTGTAGCGGACATGAGAAACAGCTGTTGGGGGATTCG 2446  
|||||  
Qy 2580 CACGCTGCTGACGCCCTGTGTAGCGGACATGAGAAACAGCTGTTGGGGGATTCG 2639  
Db 2447 GCGGACGGGCTGCTCCGCTGTTGGTGGATGATTTCTTTGGTGACACCTCACCTCAC 2506  
|||||  
Qy 2640 GCGGACGGGCTGCTCCGCTGTTGGTGGATGATTTCTTTGGTGACACCTCACCTCAC 2699  
Db 2507 CCACCGGAAACCTTCTCAGGACCTGCTCGAGGGTGTCCCTGAGTATGGCTGGCTGT 2566  
|||||  
Qy 2700 CCACCGGAAACCTTCTCAGGACCTGCTCGAGGGTGTCCGAGGTGTCCCTGAGTATGGCTGGCT 2759  
Db 2567 GAACCTGGGGAAGACAGTGGTGAATTCCTCTGTAGAAAGAGAGCCCTGGGTGGCAGGC 2626  
|||||  
Qy 2760 GAACCTGGGGAAGACAGTGGTGAATTCCTCTGTAGAAAGAGAGCCCTGGGTGGCAGGC 2819  
Db 2627 T 2627  
Qy 2820 T 2820

RESULT 14

ID US-08-951-843-100 STANDARD; DNA; UNC; 2171 BP.

AC xxxxxx

DT

DE Sequence 100, Application US/08851843  
CC Sequence 100, Application US/08851843  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Novel Telomerase  
CC NUMBER OF SEQUENCES: 223  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 100:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2171 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 22..1716  
CC OTHER INFORMATION: /note= "EcoRI-NotI insert of  
CC OTHER INFORMATION: clone 712562 encoding 63 kDa  
CC OTHER INFORMATION: telomerase protein"  
SQ SEQUENCE 2171 BP; 433 A; 567 C; 641 G; 429 T; 1 OTHER.

Query Match  
Best Local Similarity 79.6%; Pred. No. 0.00e+00;  
Matches 717; Conservative 1; Mismatches 0; Indels 183; Gaps 2;

Db 231 GTCCAGACTCCGCTTTCATCCCAAGCCTGACGGGCTGCGCGGATGTGAACATGGACTA 290  
|||  
Qy 1920 GTCCAGACTCCGCTTTCATCCCAAGCCTGACGGGCTGCGCGGATGTGAACATGGACTA 1979  
|||  
Db 291 CGTCGTGGGAGCCAGAACGTTCCGCGAGAGAAAGRGCGCTCCACCTCGAGGGT 350  
|||  
Qy 1980 CGTCGTGGGAGCCAGAACGTTCCGCGAGAGAAAGRGCGCTCCACCTCGAGGGT 2039

Db 351 GAAGCACTGTTACGCTGCTCACTAGAGGGGCGGGGCGCCCGGCTCTCTGGGCGC 410  
|||  
Qy 2040 GAAGCACTGTTACGCTGCTCACTAGAGGGGCGGGGCGCCCGGCTCTCTGGGCGC 2099  
|||  
Db 411 CTCTGTGTGGGCTTGGAGGATATCCACAGGCGCTTGGCGCACCTTCTGCTGCTGGG 470  
|||  
Qy 2100 CTCTGTGTGGGCTTGGAGGATATCCACAGGCGCTTGGCGCACCTTCTGCTGCTGGG 2159  
|||  
Db 471 GCGCCAGGACCGCGCGCTGAGCTGCTTGTCAAGTGGATGTGACGGGCGCTAGCA 530  
|||  
Qy 2160 GCGCCAGGACCGCGCGCTGAGCTGCTTGTCAAGTGGATGTGACGGGCGCTAGCA 2219  
|||  
Db 531 CACCATCCCCAGGACAGGCTACCGAGGTCATCCCGAGCATCATCAAACCCAGAACAC 590  
|||  
Qy 2220 CACCATCCCCAGGACAGGCTACCGAGGTCATCCCGAGCATCATCAAACCCAGAACAC 2279  
|||  
Db 591 GTACTGCGTGGCTGGTATGCGGTGTCAGAGGCCGCC-ATGGGCGAGCTCCGCAAGGC 649  
|||  
Qy 2280 GTACTGCGTGGCTGGTATGCGGTGTCAGAGGCCGCCCATGGGCGAGCTCCGCAAGGC 2339  
|||  
Db 650 CTTCAGAGACCGACGTC----- 665  
|||  
Qy 2340 CTTCAGAGACCGACGTCCTTACCTTCAGAGACCTCCAGCGGTACATGCGACAGTTCGTGGC 2399  
|||  
Db 665 ----- 665  
|||  
Qy 2400 TCACCTCGAGGAGACCGCGCTCAGGGATGCCGTCATCGAGAGAGCTCCTCCT 2459  
|||  
Db 665 ----- 665  
|||  
Qy 2460 GAATGAGGCCAGCAGTGGCCTTTCGACGCTTCTTCTAGCCTTCAATGTGCCACCGCGT 2519  
|||  
Db 666 -----CTAGTCCAGTGCAGGGGATCCGCGAGGCTCCATCCTCTC 707  
|||  
Qy 2520 GCGCATCAGGGCAAGTCTCTAGTCCAGTGCAGGGGATCCGCGAGGCTCATCCTCTC 2579  
|||  
Db 708 CACGCTGCTCTGACGCTGTGCTACGGCGACATGAGAACAAAGCTGTTTGGGGGATTGG 767  
|||  
Qy 2580 CACGCTGCTCTGACGCTGTGCTACGGCGACATGAGAACAAAGCTGTTTGGGGGATTGG 2639  
|||  
Db 768 GCGGACGCGGCTGCTCCCTGGTGGTGGATTTCTTGTGGTGACACCTCAGCTCAC 827  
|||  
Qy 2640 GCGGACGCGGCTGCTCCCTGGTGGTGGATTTCTTGTGGTGACACCTCAGCTCAC 2699  
|||  
Db 828 CCACGCGAAACCTTCCCTCAGGACCTGCTCGAGGCTGCTCCCTGAGTATGGCTGGTGT 887  
|||  
Qy 2700 CCACGCGAAACCTTCCCTCAGGACCTGCTCGAGGCTGCTCCCTGAGTATGGCTGGTGT 2759  
|||  
Db 888 GAACTTGGGAGAGACAGTGGTGAACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGCGC 947  
|||  
Qy 2760 GAACTTGGGAGAGACAGTGGTGAACCTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGCGC 2819  
|||  
Db 948 T 948  
Qy 2820 T 2820

RESULT 15  
ID US-08-854-050-100 STANDARD; DNA; UNC; 2171 BP.  
AC xxxxxx

DE Sequence 100, Application US/08854050  
CC Sequence 100, Application US/08854050  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Novel Telomerase

CC NUMBER OF SEQUENCES: 225  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PatentIn Release #1.0, version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 100:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2171 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 22..1716  
CC OTHER INFORMATION: /note= "EcoRI-NotI insert of  
CC OTHER INFORMATION: clone 712562 encoding 63 kDa  
CC OTHER INFORMATION: telomerase protein"  
CC SEQUENCE 2171 BP; 433 A; 567 C; 641 G; 429 T; 1 OTHER.

Query Match 52.9%; Score 477; DB 21; Length 2171;  
Best Local Similarity 79.6%; Pred. No. 0.00e+00;  
Matches 717; Conservative 1; Mismatches 0; Indels 183; Gaps 2;  
Db 231 GTCCAGACTCGGCTTCATCCCAAGCCTGACGGCTGGCGCGGATTTGTAACATGGACTA 290  
Qy 1920 GTCCAGACTCGGCTTCATCCCAAGCCTGACGGCTGGCGCGGATTTGTAACATGGACTA 1979  
Db 291 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCCGCGGCTCTCACCTCGAGGGT 350  
Qy 1980 CGTCGTGGAGCCAGAACGTTCCGACAGAAAAGAGGCCGCGGCTCTCACCTCGAGGGT 2039  
Db 351 GAAGGCACTGTTACGCGTGCTCAACTACGAGCGGCGCGGCCCGGCCCTCTCTGGGCGC 410  
Qy 2040 GAAGGCACTGTTACGCGTGCTCAACTACGAGCGGCGGCGGCCCGGCCCTCTCTGGGCGC 2099  
Db 411 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGGCGACCTTCGTGCTGGCTGGCG 470

Qy 2100 CTCTGTGCTGGCGCTGACGATATCCACAGGGCTGGCGACCTTCGTGCTGGCTGGCG 2159  
Db 471 GGCCCAAGACCCCGCCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGA 530  
Qy 2160 GGCCCAAGACCCCGCCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGA 2219  
Db 531 CACATCCCGCCAGACAGGCTCACGGAGTCTACGCCAGCATCATCAAAACCCAGACAC 590  
Qy 2220 CACATCCCGCCAGACAGGCTCACGGAGTCTACGCCAGCATCATCAAAACCCAGACAC 2279  
Db 591 GTACTGCGTGGTGGTATGCCGTGTGTCAGAAAGCGCGCC-ATGGGCGACGTCGCGCAAGGC 649  
Qy 2280 GTACTGCGTGGTGGTATGCCGTGTGTCAGAAAGCGCGCCATGGGCACGTCGCGCAAGGC 2339  
Db 650 CTTCAAGAGCCACGTC-----  
Qy 2340 CTTCAAGAGCCACGTCCTCTACTTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGC 2399  
Db 665 -----  
Qy 2400 TCACCTGACAGAGACACAGCCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCTCCCT 2459  
Db 665 -----  
Qy 2460 GAATGAGCCAGCAGTGGCGCTCTTCGACGCTTCTCTACGCTTCATGTGCCACACGCCGT 2519  
Db 666 -----CTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTC 707  
Qy 2520 GCGCATCAGGGGCAAGTCCTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTC 2579  
Db 708 CAGCTGCTCTGACGCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 767  
Qy 2580 CAGCTGCTCTGACGCTGTGTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCG 2639  
Db 768 GCGGAGCGGCTGCTCTGCTGTTGGTGATGATTTCTTGTGTGACACTCACCTCAC 827  
Qy 2640 GCGGAGCGGCTGCTCTGCTGTTGGTGATGATTTCTTGTGTGACACTCACCTCAC 2699  
Db 828 CCACGCGAAAACCTTCTCAGGACCCCTGGTCCGAGGTGTCCCTGAGTATGGTGCCTGGT 887  
Qy 2700 CCACGCGAAAACCTTCTCAGGACCCCTGGTCCGAGGTGTCCCTGAGTATGGTGCCTGGT 2759  
Db 888 GAACCTCGGGAAGACAGTGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCACGGC 947  
Qy 2760 GAACCTCGGGAAGACAGTGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGCACGGC 2819  
Db 948 T 948  
Qy 2820 T 2820

Search completed: Thu Dec 24 19:19:19 1998  
Job time : 2467 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*

W P S R L H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 18:29:00 1998; MasPar time 74.84 Seconds  
Tabular output not generated. 561.595 Million cell updates/sec

Title: >US-08-951-733-13  
Description: (1920-2820) from US08951733.seq (2 of 2)  
Perfect Score: 901  
N.A. Sequence: 1920 GTCCAGACTCGCTTCATCC.....AGGCCCTGGTGGCAGCGCT 2820  
Comp: CAGCTCTGAGCGGAAGTAGG.....TCGGGGACCCAGCGTGGCGA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 88822 seqs, 23323279 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles1

Statistics: Mean 8.416; Variance 4.601; scale 1.829

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	39	4.3	7218	1	US-08-232-Sequence 14, Applicati	4.38e-11
2	30	3.3	215	1	US-08-238-Sequence 5, Applicatio	1.52e-05
3	28	3.1	215	1	US-08-238-Sequence 5, Applicatio	2.24e-04
4	26	2.9	74	2	PCT-US95-1Sequence 94, Applicati	3.08e-03
5	26	2.9	81	2	PCT-US95-1Sequence 92, Applicati	3.08e-03
6	25	2.8	74	2	PCT-US95-1Sequence 100, Applicat	1.11e-02
7	25	2.8	75	2	PCT-US95-1Sequence 99, Applicati	1.11e-02
8	25	2.8	81	2	PCT-US95-1Sequence 98, Applicati	1.11e-02
9	25	2.8	81	2	PCT-US95-1Sequence 92, Applicati	1.11e-02
10	25	2.8	82	2	PCT-US95-1Sequence 97, Applicati	1.11e-02
11	25	2.8	7218	1	US-08-232-Sequence 14, Applicati	1.11e-02
12	24	2.7	74	2	PCT-US95-1Sequence 100, Applicat	3.90e-02
13	24	2.7	74	2	PCT-US95-1Sequence 94, Applicati	3.90e-02
14	24	2.7	75	2	PCT-US95-1Sequence 99, Applicati	3.90e-02
15	24	2.7	81	2	PCT-US95-1Sequence 98, Applicati	3.90e-02
16	24	2.7	82	2	PCT-US95-1Sequence 97, Applicati	3.90e-02
17	23	2.6	66	1	US-08-471-Sequence 144, Applicat	1.34e-01
18	23	2.6	66	1	US-08-471-Sequence 144, Applicat	1.34e-01
19	23	2.6	69	1	US-08-471-Sequence 142, Applicat	1.34e-01

C	20	23	2.6	69	1	US-08-471-Sequence 142, Applicat	1.34e-01
	21	23	2.6	84	1	US-08-209-Sequence 23, Applicati	1.34e-01
	22	23	2.6	84	1	US-08-133-Sequence 120, Applicat	1.34e-01
	23	23	2.6	84	2	PCT-US94-0Sequence 25, Applicati	1.34e-01
	24	23	2.6	84	2	PCT-US94-0Sequence 25, Applicati	1.34e-01
	25	23	2.6	84	1	US-08-300-Sequence 25, Applicati	1.34e-01
	26	23	2.6	242	1	US-08-273-Sequence 1, Applicatio	1.34e-01
C	27	22	2.4	65	1	US-08-471-Sequence 145, Applicat	4.51e-01
C	28	22	2.4	68	1	US-08-471-Sequence 143, Applicat	4.51e-01
C	29	22	2.4	2116	1	US-08-701-Sequence 1, Applicatio	4.51e-01
	30	22	2.4	7001	1	US-08-258-Sequence 1, Applicatio	4.51e-01
	31	22	2.4	7001	1	US-08-456-Sequence 1, Applicatio	4.51e-01
	32	22	2.4	7001	1	US-08-457-Sequence 1, Applicatio	4.51e-01
	33	22	2.4	7001	1	US-08-458-Sequence 1, Applicatio	4.51e-01
	34	22	2.4	7001	1	US-08-457-Sequence 1, Applicatio	4.51e-01
C	35	21	2.3	59	2	PCT-US95-1Sequence 95, Applicati	1.47e+00
	36	21	2.3	65	1	US-08-471-Sequence 145, Applicat	1.47e+00
C	37	21	2.3	66	2	PCT-US95-1Sequence 93, Applicati	1.47e+00
	38	21	2.3	68	1	US-08-471-Sequence 143, Applicat	1.47e+00
	39	21	2.3	242	1	US-08-273-Sequence 1, Applicatio	1.47e+00
	40	21	2.3	957	3	US-07-745-Sequence 16, Applicati	1.47e+00
	41	21	2.3	1387	3	5270178-1Patent No. 5270178	1.47e+00
	42	21	2.3	1755	3	5225537-1Patent No. 5225537	1.47e+00
	43	21	2.3	5467	1	US-07-745-Sequence 12, Applicati	1.47e+00
	44	21	2.3	5904	2	PCT-US92-0Sequence 1, Applicatio	1.47e+00
C	45	20	2.2	2671	3	5168051-9Patent No. 5168051	4.67e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx  
DE Sequence 14, Application US/08232463  
CC Sequence 14, Application US/08232463  
CC Patent No. 5670367  
CC GENERAL INFORMATION:  
CC APPLICANT: DORNER, F.  
CC APPLICANT: SCHEIFLINGER, F.  
CC APPLICANT: FALKNER, F. G.  
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 1800 Diagonal Road, Suite 500  
CC CITY: Alexandria  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22313-0299  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/232,463  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/935,313  
CC FILING DATE:  
CC APPLICATION NUMBER: EP 91 114 300.6  
CC FILING DATE: 26-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 30472/114 IMM  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703)836-9300  
CC TELEFAX: (703)683-4109  
CC TELEX: 899149  
CC INFORMATION FOR SEQ ID NO: 14:



```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 3.1%; Score 28; DB 1; Length 215;
Best Local Similarity 14.1%; Pred. No. 2.24e-04;
Matches 24; Conservative 62; Mismatches 84; Indels 0; Gaps 0;

Db 35 DCCNRTGVCDDTTRVYNDSSGHNKYSANYNGGNNVGAAKTHYTHTVNVSGADSKTV 94
Cp 2506 ACATGACGCTAGGAGACGCTGAGAGAGCCACTGCTGCTCATTCAGGAGGAGCTCT 2447
Db 95 TDSYNASTSSNGTGDNRSGADSYGSSKTMATSRNRTGKTANNVDSRNMGDASVGS 154
Cp 2446 GCTCATGACGAGCATCCCTCAGCGGCTGCTCTCGAGGTGAGCCAGCACTGTC 2387
Db 155 KNTKHKANSADGVGSKNNGDRNRYTGTKSVNSNCGGNGKRDVSSY 204
Cp 2386 GCATGTACGGCTGGAGGTCTCTCAAGGTAGAGACGTGCTCTTGAAGGCC 2337

RESULT 4
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DE
DT
DE Sequence 94, Application PC/TUS9511934
DE Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytozen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mirock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match 2.9%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 69
Qy 2448 GAGCTCTCCCTGATGAGCCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTG 2507
Db 70 BNNBNACGCC 79
Qy 2508 CCACCACGCC 2517

Query Match 2.9%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 69
Qy 2448 GAGCTCTCCCTGATGAGCCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTG 2507
Db 70 BNNBNACGCC 79
Qy 2508 CCACCACGCC 2517

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
```

```
Query Match 2.9%; Score 26; DB 2; Length 74;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 3 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 62
Qy 2448 GAGCTCTCCCTGATGAGCCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTG 2507
Db 63 BNNBNACGCC 72
Qy 2508 CCACCACGCC 2517

RESULT 5
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DE
DT
DE Sequence 92, Application PC/TUS9511934
DE Sequence 92, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytozen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mirock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 2.9%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 69
Qy 2448 GAGCTCTCCCTGATGAGCCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTG 2507
Db 70 BNNBNACGCC 79
Qy 2508 CCACCACGCC 2517

Query Match 2.9%; Score 26; DB 2; Length 81;
Best Local Similarity 11.4%; Pred. No. 3.08e-03;
Matches 8; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

Db 10 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 69
Qy 2448 GAGCTCTCCCTGATGAGCCAGTGGCTCTTCGACGCTTCCTACGCTTCATGTG 2507
Db 70 BNNBNACGCC 79
Qy 2508 CCACCACGCC 2517

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
```









[illegible]

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*  
M P E R L H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Thu Dec 24 19:23:48 1998; MasPar time 12290.10 Seconds  
681.038 Million cell updates/sec

Tabular output not generated.

Title: >US-08-951-733-19  
Description: (1-3798) from US08951733.seq  
Perfect Score: 3798  
N.A. Sequence: 1 CCACGGCTCCGGCAGCGCT.....GGAATAGTCCATCCCTGAT 3798  
Comp: GGTGGCAGGCGCCGTCGCA.....CCTTATCAGGTAGGGACAT

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb155  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_v1

Database: genbank107  
15:gb\_ba1 16:gb\_ba2 17:gb\_htg 18:gb\_in 19:gb\_om 20:gb\_ov  
21:gb\_pat 22:gb\_ph 23:gb\_pl1 24:gb\_pl2 25:gb\_pr1  
26:gb\_pr2 27:gb\_pr3 28:gb\_ro 29:gb\_st 30:gb\_sts 31:gb\_sy  
32:gb\_un 33:gb\_v1

Statistics: Mean 11.964; Variance 6.048; scale 1.978

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Pred. No.
1	3784	99.6	4015
2	3770	99.3	4027
3	3439	90.5	8960
4	1079	28.4	3369
5	1079	28.4	3426
6	129	3.4	201
7	67	1.8	7218
8	49	1.3	7218
9	37	1.0	74371
10	34	0.9	215
11	35	0.9	215
12	34	0.9	74371

13	30	0.8	201	21	A10161	Synthetic DNA for prep	5.30e-02
14	30	0.8	201	21	A10158	S. griseus gene for pre	5.30e-02
15	30	0.8	201	21	A10162	Synthetic DNA for prep	5.30e-02
16	30	0.8	201	21	A10159	S. griseus gene for pre	5.30e-02
17	30	0.8	201	21	E04076	gDNA encoding envelope	5.30e-02
18	29	0.8	6011	33	HS4FSIRZ2T	Epstein-Barr virus (wi	1.81e-01
19	29	0.8	15807	15	MEU72662	Methylobacterium extor	1.81e-01
20	29	0.8	76094	25	HS179D3A	Human DNA sequence fro	1.81e-01
21	32	0.8	92879	17	AC004450	*** SEQUENCING IN PROG	4.23e-03
22	31	0.8	216021	26	HUAC004787	Homo sapiens Chromosom	1.51e-02
23	23	0.7	565	21	E04076	gDNA encoding envelope	1.96e+00
24	27	0.7	904	33	HS4BHLF	Epstein-Barr virus Bam	1.96e+00
25	27	0.7	1572	19	BTU21091	Bos taurus DNA polymer	1.96e+00
26	28	0.7	1663	19	MVU92534	Mustela vison microsat	6.04e-01
27	26	0.7	2926	15	SMPROBAG	S. marcescens proBA gen	6.19e+00
28	27	0.7	4053	18	DRO28DC162	Drosophila melanogaste	1.96e+00
29	27	0.7	5050	33	HS4U2IR2A	Epstein-Barr virus (B9	1.96e+00
30	27	0.7	5400	15	SCACTVA	S. coelicolor 6 actVA r	1.96e+00
31	26	0.7	5591	15	AF027439	Acetobacter vinelandii	6.19e+00
32	28	0.7	31812	24	SPUNK4	S. pombe chromosome I c	6.04e-01
33	27	0.7	35980	16	MTY20H10	Mycobacterium tubercul	1.96e+00
34	28	0.7	39073	23	SPAC2E11	S. pombe chromosome I c	6.04e-01
35	26	0.7	43034	26	HSN5H6	Human DNA sequence fro	6.19e+00
36	27	0.7	64078	18	AC004365	Drosophila melanogaste	1.96e+00
37	27	0.7	81023	18	AC001659	Drosophila melanogaste	1.96e+00
38	26	0.7	83957	27	HS885E17	Human DNA sequence ***	1.96e+00
39	26	0.7	85132	26	U82668	Homo sapiens shox gene	6.19e+00
40	27	0.7	92879	17	AC004450	*** SEQUENCING IN PROG	1.96e+00
41	26	0.7	138706	17	AC004946	*** SEQUENCING IN PROG	6.19e+00
42	26	0.7	147542	17	HS754E20	Human DNA sequence ***	6.19e+00
43	27	0.7	172281	33	EBV	Epstein-Barr virus (EB	1.96e+00
44	27	0.7	184113	33	HS4B958RAJ	Epstein-Barr virus, ar	1.96e+00
45	27	0.7	216021	26	HUAC004787	Homo sapiens Chromosom	1.96e+00

ALIGNMENTS

RESULT	1	AF015950	4015 bp	16-AUG-1997
LOCUS		Homo sapiens telomerase reverse transcriptase (hTERT) mRNA, complete cds.		
DEFINITION				
ACCESSION		AF015950		
NID		92330016		
KEYWORDS		human.		
SOURCE		Homo sapiens		
ORGANISM		Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1	(bases 1 to 4015)		
AUTHORS		Nakamura, T.M., Morin, G.B., Chapman, K.B., Weinrich, S.L., Andrews, W.H., Lingner, J., Harley, C.B. and Cech, T.R.		
TITLE		Telomerase catalytic subunit homologs from fission yeast and human		
JOURNAL		Science 277 (5328), 955-959 (1997)		
MEDLINE		97400623		
REFERENCE	2	(bases 1 to 4015)		
AUTHORS		Morin, G.B.		
TITLE		Direct Submission		
JOURNAL		Submitted (24-JUL-1997) Geron Corporation, 230 Constitution Drive, Menlo Park, CA 94025, USA		
FEATURES		Location/Qualifiers		
source		1..4015		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/tissue_type="kidney"		
		/map="5p"		
		/dev_stage="embryo"		
		/chromosome="5"		
gene		1..4015		
		/gene="hTERT"		
CDS		56..3454		
		/gene="hTERT"		
		/codon_start=1		

/product="telomerase reverse transcriptase"

/db\_xref="pid:g3330017"

/translation="MPRAPRCRAVRSLRRSHYREVLPATFVRRLLPGQWRLLVQRGDP  
AAFLRALVQCLVCPWDAPPPAPSFQVQSLKELVARVLRQCLERGAKNVLAFGFA  
LIDGARGGPEAFVTSVSYLPNTVDALRGSGAWGLLLRRVGGDVLVHLARCALFV  
LVAPSCAYOVCPPLVOLGAATAORPPPHASGPRRRIGCCERAWNHSHVREAGVPLGLPA  
PGARRGSSASRSLPLPKPRRRCGAPEPERTVQGSWAHPGRTGRPSDRGRVCVSPSA  
RPAEATISLEBALSUSGRHSHPSVGRQHAGPPSTRPPRWDTPCPFPVIAETKHFLLYS  
SGDKQLRFSFLUSSRLTSGARLVETIFLGRSPMEPTPRRLPLPQRYWOMRPL  
FLLEGNHQAQCYGVLLKTHCPRLAAVTPAAGAREKPGQSVAAPEEDTDPRLVQ  
LLRHQSSPMOYGFVRCALRLRPVPLGLWGRSHNRERFLNKKFISLGHAKHLSLOEL  
TKMSVRDCAWLRRSPGVCPAAEHLREELIAKFLHMLSVVYVELLARSFYVET  
TFQKNRLFFYKSWNSKLASIGIRQHLKRVQLRELSEAEVRQHRARPALLSRARFI  
PKFDGLRPIVNMIDYVGRFIREKRAERLTSRKALFVSLNIRERARFGLLGASVIG  
LDDIHAWRTFVLVRQAQDPPPELFYFVDVDTGAYDIFPODLREYVIAIIIPONTY  
VRYAVVQAAHGHVAKAFKSHVSTLDLOPMRQFVAHLQETSPLRDAVVIQSSSL  
NEASSGLFDVFMCHHVRIRKSVYOCQPOGSLSTLCLSYCDGMENKLPAG  
IRRDGLLLDLVDLVLPHLTHAKTFLRLVGRVPEYGCVNLRKTVNFPVEDEAL  
GGTAFVQMPAHGLFPFCGLLLDTRILEVQSDYSYARTSIRASLTNRGFKAGRNMR  
KLFGLVRLKCHSLFDLQVNSLQTVCTNIILKLLQYRHFACVLOLPHFQVWKNPT  
FLRVISDTSALCYSLKAKNAGMSLGAAGPLPSEAVOMLCHQAFLLKLTRHRT  
YVPLGLSLTAOTQLSRKLPGLTTLTALAAANPALPSDFKTLID"

BASE COUNT 663 a 1363 c 1275 g 714 t

Query Match 99.6%; Score 3784; DB 27; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 3785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db	1	GCACGCTGGTCTCTGCTGCGCAGCTGGGAGCCCTGGCCCCGCGCACCCCGCGATGCC	60
QY	13	GCACGCTGGTCTCTGCTGCGCAGCTGGGAGCCCTGGCCCCGCGCACCCCGCGATGCC	72
Db	61	GCAGCGTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGAGCCACTACCGCGAGGTGCT	120
QY	73	GCAGCGTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGAGCCACTACCGCGAGGTGCT	132
Db	121	GCCGCTGGCAGCTTGTGGGGCCCTGGGGCCCGCCAGGCTGGGGCTGTGTGACGCGGG	180
QY	133	GCCGCTGGCAGCTTGTGGGGCCCTGGGGCCCGCCAGGCTGGGGCTGTGTGACGCGGG	192
Db	181	GGACCCGCGGCTTTCGCGCGCTGTGGGCCAGTGGTGGTGGTGGTGGTGGTGGGAGCG	240
QY	193	GGACCCGCGGCTTTCGCGCGCTGTGGGCCAGTGGTGGTGGTGGTGGTGGTGGGAGCG	252
Db	241	ACGCGCCGCCCCCGCCCTCTTCCGCGCAGGTGTCTGCTGAGGAGGTGTGTGGTGGC	300
QY	253	ACGCGCCGCCCCCGCCCTCTTCCGCGCAGGTGTCTGCTGAGGAGGTGTGTGGTGGC	312
Db	301	CCGAGTGTGAGAGGTGTGCGAGCGGGCGCGGAGAAAGTGTGTGCGCTTGGGCTTCGC	360
QY	313	CCGAGTGTGAGAGGTGTGCGAGCGGGCGCGGAGAAAGTGTGTGCGCTTGGGCTTCGC	372
Db	361	GCTCTGAGCGGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA	420
QY	373	GCTCTGAGCGGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA	432
Db	421	CCTCCCAACACGCTGACCGACGCTGCGGGGAGCGGGCGGTGCTGCTGCGG	480
QY	433	CCTCCCAACACGCTGACCGACGCTGCGGGGAGCGGGCGGTGCTGCTGCGG	492
Db	481	CCGCGTGGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540
QY	493	CCGCGTGGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	552
Db	541	GGCTCCGAGCTGCGCTTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG	600
QY	553	GGCTCCGAGCTGCGCTTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG	612
Db	601	TCAGGCG	660
QY	613	TCAGGCG	672

Db	661	CTGGAACCATAGCGTCAGGAGGCCCGGGTCCCGCTGGCGCTGCCAGCCCCCGGGTGGGAG	720
QY	673	CTGGAACCATAGCGTCAGGAGGCCCGGGTCCCGCTGGCGCTGCCAGCCCCCGGGTGGGAG	732
Db	721	GAGCGCGGGGCGAGTCCAGCGGAAGTCTGCGGTTGGCCCAAGAGCCGACAGGCGTGGCGC	780
QY	733	GAGCGCGGGGCGAGTCCAGCGGAAGTCTGCGGTTGGCCCAAGAGCCGACAGGCGTGGCGC	792
Db	781	TGCCCCCTAGCGCGGAGCGGACGCCCTTGGGCGAGGGTCTCTGGGCGCCACCCCGGCGAGGAC	840
QY	793	TGCCCCCTAGCGCGGAGCGGACGCCCTTGGGCGAGGGTCTCTGGGCGCCACCCCGGCGAGGAC	852
Db	841	GCCTGACCGAGTGACCGTGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
QY	853	GCCTGACCGAGTGACCGTGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	912
Db	901	CACCTCTTTGGAGGGTGGCTCTCTGGCAGCGGCGCACTCCCAACCCATCGTGGGCGCGCA	960
QY	913	CACCTCTTTGGAGGGTGGCTCTCTGGCAGCGGCGCACTCCCAACCCATCGTGGGCGCGCA	972
Db	961	GCACACGCGGGCGCCCCCATCCACATCGGGGCGCACACCGTCCCTGGGACACGCGTTGTCC	1020
QY	973	GCACACGCGGGCGCCCCCATCCACATCGGGGCGCACACCGTCCCTGGGACACGCGTTGTCC	1032
Db	1021	CCCGGTGTAGCGCGAGCAACAGCACTTCTCTACTCTCAGGCGCAAGAGAGCAGTGGC	1080
QY	1033	CCCGGTGTAGCGCGAGCAACAGCACTTCTCTACTCTCAGGCGCAAGAGAGCAGTGGC	1092
Db	1081	GCCTCTTCTCTACTCTGAGGCGCCAGCCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCT	1140
QY	1093	GCCTCTTCTCTACTCTGAGGCGCCAGCCTGAGGCGCTGAGGCGCTGAGGCGCTGAGGCGCT	1152
Db	1141	GACCATCTTCTGGGTTCAGGCGCTTGGATGCCAGGAGTCCCGCGAGGTTGGCCGCGCT	1200
QY	1153	GACCATCTTCTGGGTTCAGGCGCTTGGATGCCAGGAGTCCCGCGAGGTTGGCCGCGCT	1212
Db	1201	GCCCCAGCGCTACTGCGAAATGCGGCCCTGTCTTGTGGAGTCTTGGGAAACACACGCGCA	1260
QY	1213	GCCCCAGCGCTACTGCGAAATGCGGCCCTGTCTTGTGGAGTCTTGGGAAACACACGCGCA	1272
Db	1261	GTCCCTCTAGGGGTGCTCTCAAGACGACATGCCCGCTGCGAGCTGCGGTTCACCCGAGC	1320
QY	1273	GTCCCTCTAGGGGTGCTCTCTCAAGACGACATGCCCGCTGCGAGCTGCGGTTCACCCGAGC	1332
Db	1321	AGCGGTGTCTGTGCGCGGAGAACCCAGGCTCTGTGGCGGCCCGCGAGGAGGAGGA	1380
QY	1333	AGCGGTGTCTGTGCGCGGAGAACCCAGGCTCTGTGGCGGCCCGCGAGGAGGAGGA	1392
Db	1381	CACAGACCCCGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1440
QY	1393	CACAGACCCCGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1452
Db	1441	CGGCTTGTGCGGCGCTGCTGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1500
QY	1453	CGGCTTGTGCGGCGCTGCTGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1512
Db	1501	CAACGAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAA	1560
QY	1513	CAACGAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAA	1572
Db	1561	GCTCTCGCTGACGAGCTGACGTGGAGATGAGCTGCGGGAGTGCCTGCGTGGCGCAG	1620
QY	1573	GCTCTCGCTGACGAGCTGACGTGGAGATGAGCTGCGGGAGTGCCTGCGTGGCGCAG	1632
Db	1621	GAGCCAGGGGTGGCTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1680
QY	1633	GAGCCAGGGGTGGCTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1692
Db	1681	CAAGTTCCTGCACTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTCTTCTTTTA	1740
QY	1693	CAAGTTCCTGCACTGGCTGATGAGTGTACGTGCTGAGCTGCTCAGGCTCTTCTTTTA	1752
Db	1741	TGTCACGAGACCAACGCTTTCAAAGAACAGGCTCTTTTCTTACCGGAAGAGTGTCTGGAG	1800



[illegible]



QY 1223 TACTGGCAATCGGGCCCTGTTCTGGAGCTGCTTGGGAACACCGCGAGTGCCTTAC 1282  
Db 1272 GGGGTGCTCCTCAAGACGACATGCCGCTGCGAGTGCCTGTCACCCACAGCCGGTGTG 1331  
QY 1283 GGGGTGCTCCTCAAGACGACATGCCGCTGCGAGTGCCTGTCACCCACAGCCGGTGTG 1342  
Db 1332 TGTGCCGGGAAGCCCGAGGGCTGTGTGGGGCCCCCGAGGAGGAGACACAGACCCC 1391  
QY 1343 TGTGCCGGGAAGCCCGAGGGCTGTGTGGGGCCCCCGAGGAGGAGACACAGACCCC 1402  
Db 1392 CGTGGCTGTGACGTGCTCGCCACAGACAGCAGACCCCTGCGAGGTGTACGGCTTCGTG 1451  
QY 1403 CGTGGCTGTGACGTGCTCGCCACAGACAGACGCCCTGCGAGGTGTACGGCTTCGTG 1462  
Db 1452 CGGGCTGTGCTCGCGGGGTGTGTCCCCACAGGCTCTGGGGCTTCCAGGACACAGAACCC 1511  
QY 1463 CGGGCTGTGCTCGCGGGGTGTGTCCCCACAGGCTCTGGGGCTTCCAGGACACAGAACCC 1522  
Db 1512 CGTTCCTCAGGAACACCAAGATTCATCTCCTGGGGAAGCATGCGCAAGCTTCGCTG 1571  
QY 1523 CGCTTCCTCAGGAACACCAAGATTCATCTCCTGGGGAAGCATGCGCAAGCTTCGCTG 1582  
Db 1572 CAGGAGCTCAGCTGGAAGATGAGCTGCGGGGTGCGCTGCTGCGAGGAGCCACAGG 1631  
QY 1583 CAGGAGCTCAGCTGGAAGATGAGCTGCGGGGTGCGCTGCTGCGAGGAGCCACAGG 1642  
Db 1632 GTTGGCTGTGTTCCGGCCCGCAGACACCTCTGCTGAGGAGATCTTGCCCAAGTTCCTG 1691  
QY 1643 GTTGGCTGTGTTCCGGCCCGCAGACACCTCTGCTGAGGAGATCTTGCCCAAGTTCCTG 1702  
Db 1692 CACTGGCTGATGAGTGTACGTGTCGAGCTGCTCAGCTCTTTCTTTATGTCAGGAG 1751  
QY 1703 CACTGGCTGATGAGTGTACGTGTCGAGCTGCTCAGCTCTTTCTTTATGTCAGGAG 1762  
Db 1752 ACCACGTTTCAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTCGA 1811  
QY 1763 ACCACGTTTCAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTCGA 1822  
Db 1812 AGCATTGGAATCAGACAGACACTTGAAGAGGTGCAGCTCGGGAGCTGTGCGGAAGCAGAG 1871  
QY 1823 AGCATTGGAATCAGACAGACACTTGAAGAGGTGCAGCTCGGGAGCTGTGCGGAAGCAGAG 1882  
Db 1872 GTCAGCAGCATCGGGAAGCCAGGCCCGCTGCTGACGTGCAGACTCCGCTTCATCCCC 1931  
QY 1883 GTCAGCAGCATCGGGAAGCCAGGCCCGCTGCTGACGTGCAGACTCCGCTTCATCCCC 1942  
Db 1932 AAGCCTGACGGCTGCGGCCGATGTGAACATGGACTAGTCTGTGGAGCCAGAGGTTTC 1991  
QY 1943 AAGCCTGACGGCTGCGGCCGATGTGAACATGGACTAGTCTGTGGAGCCAGAGGTTTC 2002  
Db 1992 CGCAGAGAAAAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTACGCGTGTCTC 2051  
QY 2003 CGCAGAGAAAAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTACGCGTGTCTC 2062  
Db 2052 AACTACAGGGCGCGGGCCCGCCCGCTGCTGTGGGGCCCTCTGTGTGGCCCTGGACGAT 2111  
QY 2063 AACTACAGGGCGCGGGCCCGCCCGCTGCTGTGGGGCCCTCTGTGTGGCCCTGGACGAT 2122  
Db 2112 ATCCAGAGCCCTGCGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2171  
QY 2123 ATCCAGAGCCCTGCGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2182  
Db 2172 CTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCCGCAGGACAGGCTC 2231  
QY 2183 CTGTACTTTGTCAAGGTGGATGTGACGGCGCGGTACGACACCATCCCGCAGGACAGGCTC 2242  
Db 2232 ACGGAGGTATGCCAGCATCATCAAAACCCAGAACACGTACTGCTGCTGCTGCTGCTGCTGCTG 2291  
QY 2243 ACGGAGGTATGCCAGCATCATCAAAACCCAGAACACGTACTGCTGCTGCTGCTGCTGCTGCTG 2302  
Db 2292 GTGGTCCAGNAGCCCGCCATGCGGACGTCCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2351  
QY 2303 GTGGTCCAGNAGCCCGCCATGCGGACGTCCGCAAGGCTTCAAGAGCCAGCTCTCTACC 2362

Db 2352 TTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGTCTACCTTCAGAGGAGACACGCCG 2411  
QY 2363 TTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGTCTACCTTCAGAGGAGACACGCCG 2422  
Db 2412 CTGAGGGATGCGGTGCTCATCGAGCAGAGTCTCTCCCTGAATGAGGCCAGAGTGGCTC 2471  
QY 2423 CTGAGGGATGCGGTGCTCATCGAGCAGAGTCTCTCCCTGAATGAGGCCAGAGTGGCTC 2482  
Db 2472 TTGAGAGCTTCTCCTACGCTTCATGTGCCACACAGCGCTGCGCATCAGGGGCAAGTCCTAC 2531  
QY 2483 TTGAGAGCTTCTCCTACGCTTCATGTGCCACACAGCGCTGCGCATCAGGGGCAAGTCCTAC 2542  
Db 2532 GTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTCTCTGACGCTGTGC 2591  
QY 2543 GTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTCTCTGACGCTGTGC 2602  
Db 2592 TAGCGGCACATGGAAGAACAGCTGTTGCGGGGATTCGGCGGAGTGGCTGCTCTGCT 2651  
QY 2603 TAGCGGCACATGGAAGAACAGCTGTTGCGGGGATTCGGCGGAGTGGCTGCTCTGCT 2662  
Db 2652 TTGTTGGATGATTCTTGTGTTGGTACACTCACCTACCCACGCGAAACCTTCTCTCAGG 2711  
QY 2663 TTGTTGGATGATTCTTGTGTTGGTACACTCACCTACCCACGCGAAACCTTCTCTCAGG 2722  
Db 2712 ACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2771  
QY 2723 ACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2782  
Db 2772 AACTTCCCTGTAGAAGAGAGGCCCTGGTGGCAGCGCTTTTGTTCAGATGCGCGCCAC 2831  
QY 2783 AACTTCCCTGTAGAAGAGAGGCCCTGGTGGCAGCGCTTTTGTTCAGATGCGCGCCAC 2842  
Db 2832 GSCCTATTCCCTGTGCGGCTGCTGCTGATACCCGACCTCGAGTGCAGAGCCAC 2891  
QY 2843 GSCCTATTCCCTGTGCGGCTGCTGCTGATACCCGACCTCGAGTGCAGAGCCAC 2902  
Db 2892 TACTCCAGTATGCCCGACCTCCATCAGAGCAGTCTCACCTTCAACCGCGGCTTCAAG 2951  
QY 2903 TACTCCAGTATGCCCGACCTCCATCAGAGCAGTCTCACCTTCAACCGCGGCTTCAAG 2962  
Db 2952 GCTGGAGAAACATCGTGCAGAACTCTTTGGGGTCTTGGGGCTGAAAGTGCAGAGCTG 3011  
QY 2963 GCTGGAGAAACATCGTGCAGAACTCTTTGGGGTCTTGGGGCTGAAAGTGCAGAGCTG 3022  
Db 3012 TTTCTGGATTTCACGTGAACAGCTCCAGAGCGGTGACCAACATCTACAGATCTCTC 3071  
QY 3023 TTTCTGGATTTCACGTGAACAGCTCCAGAGCGGTGACCAACATCTACAGATCTCTC 3082  
Db 3072 CTGCTGCAGGCTTACAGGTTTTCACGCATGTGTGCTGCAGCTCCCATTTTCATCAGCAAGTT 3131  
QY 3083 CTGCTGCAGGCTTACAGGTTTTCACGCATGTGTGCTGCAGCTCCCATTTTCATCAGCAAGTT 3142  
Db 3132 TGGAGAACCCACATTTTCTGCGGCTATCTCTGACAGCGGCTCTCTGCTGCTGCTGCTGCTG 3191  
QY 3143 TGGAGAACCCACATTTTCTGCGGCTATCTCTGACAGCGGCTCTCTGCTGCTGCTGCTGCTG 3202  
Db 3192 ATCTGGAAGCCAAAGCAGGATGTGCTGGGGCCAAAGGCTTCTGCTCAAGTGTACTGCAGAC 3251  
QY 3203 ATCTGGAAGCCAAAGCAGGATGTGCTGGGGCCAAAGGCTTCTGCTCAAGTGTACTGCAGAC 3262  
Db 3252 CCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTGCAGAC 3311  
QY 3263 CCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTGCAGAC 3322  
Db 3312 CGTGTACCTAGTGCAGTCTCTGGGTGCTGCTGACAGAGCCACAGCAGCTGAGTGG 3371  
QY 3323 CGTGTACCTAGTGCAGTCTCTGGGTGCTGCTGACAGAGCCACAGCAGCTGAGTGG 3382  
Db 3372 AAGCTCCCGGGACAGCTGACTGCCCTGGAGCGCGAGCCAAACCCGCGCTGCTGCTGCTGCTG 3431  
QY 3383 AAGCTCCCGGGACAGCTGACTGCCCTGGAGCGCGAGCCAAACCCGCGCTGCTGCTGCTGCTGCTG 3442



Db 1701 GCCCGCCACACCTAGTGGACCCGAAAGCGCTCTGGGATGCCAAGGGCGCTGGAACCA 1760  
|  
|  
|  
QY 622 GCCCGCCACACCTAGTGGACCCGAAAGCGCTCTGGGATGCCAAGGGCGCTGGAACCA 681  
|  
|  
|  
Db 1761 TAGCGTCAAGGAGCGCGGGTCCCTTGGGCTGCCAGCCCGGGTGGAGGAGCGCGG 1820  
|  
|  
|  
QY 682 TAGCGTCAAGGAGCGCGGGTCCCTTGGGCTGCCAGCCCGGGTGGAGGAGCGCGG 741  
|  
|  
|  
Db 1821 GGGCAGTCCAGCGGAAGTCTGCCGTTGCCAAGAGCCCAAGGCGTGGCGTGCCTCTGA 1880  
|  
|  
|  
QY 742 GGGCAGTCCAGCGGAAGTCTGCCGTTGCCAAGAGCCCAAGGCGTGGCGTGCCTCTGA 801  
|  
|  
|  
Db 1881 GCCGAGCGGACGCCCTTGGGAGGGGTCTGGGCGCCACCCGGGCGAGGAGCGGTGACC 1940  
|  
|  
|  
QY 802 GCCGAGCGGACGCCCTTGGGAGGGGTCTGGGCGCCACCCGGGCGAGGAGCGGTGACC 861  
|  
|  
|  
Db 1941 GAGTGACCGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2000  
|  
|  
|  
QY 862 GAGTGACCGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 921  
|  
|  
|  
Db 2001 GGAGGTGCGCTCTCTGT 2060  
|  
|  
|  
QY 922 GGAGGTGCGCTCTCTGT 981  
|  
|  
|  
Db 2061 GGGCCCCCATCCACATCGGGGCCACACAGTCTCTGTGTGTGTGTGTGTGTGTGTGT 2120  
|  
|  
|  
QY 982 GGGCCCCCATCCACATCGGGGCCACACAGTCTCTGTGTGTGTGTGTGTGTGTGTGT 1041  
|  
|  
|  
Db 2121 CGCGAGACCAAGCACCTCTCTACTCTCTAGGCGGACAAAGGAGAGTGTGGGCGCTCT 2180  
|  
|  
|  
QY 1042 CGCGAGACCAAGCACCTCTCTACTCTCTAGGCGGACAAAGGAGAGTGTGGGCGCTCT 1101  
|  
|  
|  
Db 2181 CCTACTCAGTCTCTGAGGCCACGCTCTGAGTGTGGGCTCGAGGCTCTGTGGAGACCATCT 2240  
|  
|  
|  
QY 1102 CCTACTCAGTCTCTGAGGCCACGCTCTGAGTGTGGGCTCGAGGCTCTGTGGAGACCATCT 1161  
|  
|  
|  
Db 2241 TCTGGGTTCCAGGCCCTTGAGTCCAGGAGTCTCCCGCAGAGTGTCCCGCTGTCCCGCAGCG 2300  
|  
|  
|  
QY 1162 TCTGGGTTCCAGGCCCTTGAGTCCAGGAGTCTCCCGCAGAGTGTCCCGCAGCGTGTCCCGCAGCG 1221  
|  
|  
|  
Db 2301 CTACTGCAATCGGCGCCCTGTTCTTCTGAGCTGCTTGGAGCCACCGCGAGTGCCTTA 2360  
|  
|  
|  
QY 1222 CTACTGCAATCGGCGCCCTGTTCTTCTGAGCTGCTTGGAGCCACCGCGAGTGCCTTA 1281  
|  
|  
|  
Db 2361 CGGGGTCTCTCTCAAGACGCACTCCCGCTCGAGCTGCGGTACCCAGCAGCGCGTGT 2420  
|  
|  
|  
QY 1282 CGGGGTCTCTCTCAAGACGCACTCCCGCTCGAGCTGCGGTACCCAGCAGCGCGTGT 1341  
|  
|  
|  
Db 2421 CTGTGCCGGAGAACCCAGAGGCTCTGTGGCGGCCCGAGGAGGAGGACAGACCC 2480  
|  
|  
|  
QY 1342 CTGTGCCGGAGAACCCAGAGGCTCTGTGGCGGCCCGAGGAGGAGGACAGACCC 1401  
|  
|  
|  
Db 2481 CCCTCGGCTGTGTGAGTGTCTCCGACGACAGACAGCCCTGGCAGTGTACGGCTTCGT 2540  
|  
|  
|  
QY 1402 CCCTCGGCTGTGTGAGTGTCTCCGACGACAGACAGCCCTGGCAGTGTACGGCTTCGT 1461  
|  
|  
|  
Db 2541 GCGGGCTGTCTGCGCGGCTGTGTGCGCCCGAGGCTCTGGGGTCCAGGACACAGAACG 2600  
|  
|  
|  
QY 1462 GCGGGCTGTCTGCGCGGCTGTGTGCGCCCGAGGCTCTGGGGTCCAGGACACAGAACG 1521  
|  
|  
|  
Db 2601 CCCTCTCTCAGGAACACCAAGATTCATCTCCCTGGGGAAGCATGCCAGCTCTCGCT 2660  
|  
|  
|  
QY 1522 CCCTCTCTCAGGAACACCAAGATTCATCTCCCTGGGGAAGCATGCCAGCTCTCGCT 1581  
|  
|  
|  
Db 2661 GCAGGAGTGCAGTGGGAAGATGAGCTGGGGGTGCGCTTGGCTGGCGAGGAGCCAGG 2720  
|  
|  
|  
QY 1582 GCAGGAGTGCAGTGGGAAGATGAGCTGGGGGTGCGCTTGGCTGGCGAGGAGCCAGG 1641  
|  
|  
|  
Db 2721 GGTGGCTGTCTCCGGCCGACAGACCGTCTCGGTGAGGAGATCTCGCCAGTTCCT 2780  
|  
|  
|  
QY 1642 GGTGGCTGTCTCCGGCCGACAGACCGTCTCGGTGAGGAGATCTCGCCAGTTCCT 1701

Db 2781 GCACTGGCTGATGAGTGTGTACGTCTCGAGCTGCTCAGGTCTTTTCTTTATGTACGGA 2840  
|  
|  
|  
QY 1702 GCACTGGCTGATGAGTGTGTACGTCTCGAGCTGCTCAGGTCTTTTCTTTATGTACGGA 1761  
|  
|  
|  
Db 2841 GACCACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGTGTCTGGAGCAAGTTGCA 2900  
|  
|  
|  
QY 1762 GACCACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGTGTCTGGAGCAAGTTGCA 1821  
|  
|  
|  
Db 2901 AAGCATTGGAATCAGACAGACATTGAAGAGGGTGCAGCTCGGGAGCTGTCTGGAAGCAGA 2960  
|  
|  
|  
QY 1822 AAGCATTGGAATCAGACAGACATTGAAGAGGGTGCAGCTCGGGAGCTGTCTGGAAGCAGA 1881  
|  
|  
|  
Db 2961 GGTGAGCAGCATCGGGAAGCCAGGCGCCCTGTGTACGTCCAGACTCCGCTTCATCCC 3020  
|  
|  
|  
QY 1882 GGTGAGCAGCATCGGGAAGCCAGGCGCCCTGTGTACGTCCAGACTCCGCTTCATCCC 1941  
|  
|  
|  
Db 3021 CAAGCCTTGACGGCTCGGCGCCGATTTGAACATGGACTAGTCTGTGGAGCCAGCAAGCTT 3080  
|  
|  
|  
QY 1942 CAAGCCTTGACGGCTCGGCGCCGATTTGAACATGGACTAGTCTGTGGAGCCAGCAAGCTT 2001  
|  
|  
|  
Db 3081 CCGCAGAGAAAAGAGGGCCGAGCGTCTCTACCTCGAGGGTGAAGGCACTGTTTCAGCGTGTCT 3140  
|  
|  
|  
QY 2002 CCGCAGAGAAAAGAGGGCCGAGCGTCTCTACCTCGAGGGTGAAGGCACTGTTTCAGCGTGTCT 2061  
|  
|  
|  
Db 3141 CAACTAGCAGCGGGCGGGCGCCCTCTCTGGGCGCTCTGTGTGGGCTGTGACGA 3200  
|  
|  
|  
QY 2062 CAACTAGCAGCGGGCGGGCGCCCTCTCTGGGCGCTCTGTGTGGGCTGTGACGA 2121  
|  
|  
|  
Db 3201 TATCCACAGGGCTGTGGCGACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3260  
|  
|  
|  
QY 2122 TATCCACAGGGCTGTGGCGACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2181  
|  
|  
|  
Db 3261 GCTGTACTTGTCAAGTGGATGTGACGGCGGTGACGACCATCCCGCAGGACAGGCT 3320  
|  
|  
|  
QY 2182 GCTGTACTTGTCAAGTGGATGTGACGGCGGTGACGACCATCCCGCAGGACAGGCT 2241  
|  
|  
|  
Db 3321 CAGGAGGTCACTCGCAGCATCATCAAAACCCAGAACACGTACTGTGGTGGCTGCGTATGC 3380  
|  
|  
|  
QY 2242 CAGGAGGTCACTCGCAGCATCATCAAAACCCAGAACACGTACTGTGGTGGCTGCGTATGC 2301  
|  
|  
|  
Db 3381 CGTGTGTCAGAAAGCGCCCATGGGCGCTCCGCAAGGCTTCAAGAGCCACGCTCTAC 3440  
|  
|  
|  
QY 2302 CGTGTGTCAGAAAGCGCCCATGGGCGCTCCGCAAGGCTTCAAGAGCCACGCTCTAC 2361  
|  
|  
|  
Db 3441 CTTGACAGACCTCCAGCGGTACATGGACAGTTCGTGGCTCACTTGAGGAGACAGCC 3500  
|  
|  
|  
QY 2362 CTTGACAGACCTCCAGCGGTACATGGACAGTTCGTGGCTCACTTGAGGAGACAGCC 2421  
|  
|  
|  
Db 3501 GCTGAGGATGCGGTCTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGAGTGGCCT 3560  
|  
|  
|  
QY 2422 GCTGAGGATGCGGTCTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGAGTGGCCT 2481  
|  
|  
|  
Db 3561 CTTGAGGCTTCTCTAGCTTCATGTGCCACAGCGCGTGGCATCAGGGGCAAGTCTTA 3620  
|  
|  
|  
QY 2482 CTTGAGGCTTCTCTAGCTTCATGTGCCACAGCGCGTGGCATCAGGGGCAAGTCTTA 2541  
|  
|  
|  
Db 3621 CGTCCAGTGGCAGGGATCCGAGGGCTCCATCTCTCCACGCTGCTCTGACAGCTGTG 2601  
|  
|  
|  
QY 2542 CGTCCAGTGGCAGGGATCCGAGGGCTCCATCTCTCCACGCTGCTCTGACAGCTGTG 2601  
|  
|  
|  
Db 3681 CTACGGGACATGGAACAAAGCTGTTGGGGGATTCGGCGGAGCGGGCTGCTCTCGG 3740  
|  
|  
|  
QY 2602 CTACGGGACATGGAACAAAGCTGTTGGGGGATTCGGCGGAGCGGGCTGCTCTCGG 2661  
|  
|  
|  
Db 3741 TTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3800  
|  
|  
|  
QY 2662 TTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2721  
|  
|  
|  
Db 3801 GACCTGGTCCGAGGTGCTCCTGAGTATGCTGCTGGTGAACCTGCGGAGACAGTGGT 3860  
|  
|  
|  
QY 2722 GACCTGGTCCGAGGTGCTCCTGAGTATGCTGCTGGTGAACCTGCGGAGACAGTGGT 2781  
|  
|  
|  
Db 3861 GAACTTCCCTGTAGAAGACGAGGCCCTGGTGGCAGGGCTTTTGTTCAGATGCGGCCCA 3920

```
Qy 2782 GAACCTCCCTGTAGAACGACGAGCCCTGGGTGGACAGCGCTTTGTTGATGCGCGGCCA 2841
Db 3921 CGGCCTATTCCTCCCTGGTGGCGCTGCTGCTGGATACCCGCGCCCTGGAGGTGCGAGCGGA 3980
Qy 2842 CGGCCTATTCCTCCCTGGTGGCGCTGCTGCTGGATACCCGCGCCCTGGAGGTGCGAGCGGA 2901
Db 3981 CTACTCCAGCTATGCCCCGAGCTCATCAGAGCCAGTCTCAGCTTCAACCGCGGCTTCAA 4040
Qy 2902 CTACTCCAGCTATGCCCCGAGCTCATCAGAGCCAGTCTCAGCTTCAACCGCGGCTTCAA 2961
Db 4041 GCTGGGAGGACATGCGCTCGCAACTCTTGGGCTCTTGGGCTTGGGCTGAAGTGTCCAGCGCT 4100
Qy 2962 GCTGGGAGGACATGCGCTCGCAACTCTTGGGCTCTTGGGCTTGGGCTGAAGTGTCCAGCGCT 3021
Db 4101 GTTCTGGATTGCAAGTGAACAGCTCCAGAGCGTGTGCAACCACTCTACAAGATCCT 4160
Qy 3022 GTTCTGGATTGCAAGTGAACAGCTCCAGAGCGTGTGCAACCACTCTACAAGATCCT 3081
Db 4161 CTTGCTCAGGGGTACAGGTTTACAGCATGTGTGCTGAGCTCCCATTTATCATCAGCAAGT 4220
Qy 3082 CTTGCTCAGGGGTACAGGTTTACAGCATGTGTGCTGAGCTCCCATTTATCATCAGCAAGT 3141
Db 4221 TTGGAAGAACCCACATTTTCTCGCGTCTCATCTCTGACAGCGCTCCCTCTGCTACTC 4280
Qy 3142 TTGGAAGAACCCACATTTTCTCGCGTCTCATCTCTGACAGCGCTCCCTCTGCTACTC 3201
Db 4281 CATCTCTGAAGCAAGCAAGCAGGATGCTGCTGGGGCCCAAGGCGCGCGGCGCTCT 4340
Qy 3202 CATCTCTGAAGCAAGCAAGCAGGATGCTGCTGGGGCCCAAGGCGCGCGGCGCTCT 3261
Db 4341 GCCCTCCAGGGCGTACAGTGTGCTGCGCACAAGCATTTCTGCTCAAGCTGACTCGACA 4400
Qy 3262 GCCCTCCAGGGCGTACAGTGTGCTGCGCACAAGCATTTCTGCTCAAGCTGACTCGACA 3321
Db 4401 CGGTGTACCTACGTCGACCTCTCGGCTCACTAGGACAGCCAGAGCGAGCTGAGTCG 4460
Qy 3322 CGGTGTACCTACGTCGACCTCTCGGCTCACTAGGACAGCCAGAGCGAGCTGAGTCG 3381
Db 4461 GAAGTCCCGGGGAGCAGCTGACTGCTGAGGCGCCAGCCCAACCGGACCTGCCCCC 4520
Qy 3382 GAAGTCCCGGGGAGCAGCTGACTGCTGAGGCGCCAGCCCAACCGGACCTGCCCCC 3441
Db 4521 AGACTTCAAGACCATCTCTGGACT 4543
Qy 3442 AGACTTCAAGACCATCTCTGGACT 3464

RESULT 4
LOCUS AF073311 3369 bp mRNA ROD 09-SEP-1998
DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.
ACCESSION AF073311
NID g3551846
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3369)
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
TITLE Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.
TITLE Expression of mouse telomerase catalytic subunit in embryos and
adult tissues
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)
MEDLINE 98393668
REFERENCE 2 (bases 1 to 3369)
AUTHORS Martin-Rivera, L., Herrera, E. and Blasco, M. A.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) Immunology and Oncology, National Centre of
Biotechnology, Cantoblanco, Madrid 28049, Spain
FEATURES
1. 3369
source /organism="Mus musculus"
/db_xref="taxon:10090"
```

```
CDS
1..3369
/codon_start=1
/product="telomerase catalytic subunit"
/db_xref="PID:g3551847"
/translation="MTRAPRCPAVRSLRSRYREVWPLATFVRRLGPEGRRLVOPGDP
KLYRTVAOCLVCMHWSQPPADLSFHOVSILKELVAVRVORLCERNERNVLAFGFE
LNEARGGPPMAFTSSVRSYSLNTVETLRVSGANWMLLSRVGDDLLYLLHACALYL
LVPPSCAYQVCGSPLYICATDDIWDVPSASRYRTPRPVGRNFTNLRLFLQOIKSSRQE
ARSPALPSRGTRKHLSTSVSAKARCPVPRVEEGPHRQVLPSTPGSKSWPSP
AKSPVPTAEKDLSSKGVSCCKKPKPSTSLSPRQNAFOLRPFIEFRH
FLYSRDGOERLNPSFLNLOPLNTGARRVEIIFLGSRPRTSGPLCRLHRLSSPWQ
QMRPLFOQLVNHAEQVYRLLRSHCFRTANOVTDALNTPPHMLDILLHSSPWQ
VYGLFACLCCKVVSASLWGTNRNRRFFNKKFISLGKYGKLSLQELMWMKVECH
WURSSPGKDRVPAARHRLERILATFLWMDTYVVQLRSFFYITESTFKNRLFFY
KUSWSKLSQISIGVRLRERLRELSQEERHHQDTWLAMPICRLRFIPKNGRLIV
NYSYSGTRALGRKQAOHTQRLKTLFSLNAYERTKHPHLMGSSVLGNDIYRTVRA
FYLVRALDQTPMYEVKADVTGAYDAIPQGLVEVANNIRHSESTYCIROYAVVRR
DSOGOVHKSFRRQVTTLSQLPYMGOFKHLQDSASALRNSVVIQSSISMNESSSL
FFDFLHRLHVVVIGDRCYTCQGGIPQSSSLTSLCSLCFGDMENKLEAEVQRDQYL
LRVDDFLVTPHLDOAKTFLTLVHGVPEYGCMLNKTQVNVNFPVPEPTGLGAAPIYR
LPAHCLFPMWGLLDTQTLVFCDSYGAQTSIKTSLTFSQVFKAGTKMRNKLISYLR
LKCHGLFLQVNSLTQVCINIKYIFLQAYRFHACVIOLPDQVRKMLTFLGLIIS
SOASCYAILKVNKPMGTLKAGSPPEAAHWLCYQAFLLKLAHSHVYIKLLGLPLRT
AOKLCKLPKPEATMTILKAAADPALSTDFQTLTD"
BASE COUNT 737 a 959 c 891 g 782 t
ORIGIN
Query Match 28.4%; Score 1079; DB 28; Length 3369;
Best Local Similarity 68.2%; Pred. No. 0.00e+00;
Matches 2343; Conservative 0; Mismatches 988; Indels 102; Gaps 6;
Db 1 ATGACCGCGCTCTCTGTCGCCCGCGGTGGCTCTCTGTGCGCAGCGCATACCGGGAG 60
Qy 68 ATGCGCGCGCTCCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCATACCGGAG 127
Db 61 GTGTGGCGCGTGGCAACTTTGTGCGCGCGCTTGGGCGCGCGAGCGCGCTGTGTCAA 120
Qy 128 GTGTGCGCGTGGCGCGCTGCTGCGCGCGCTGGGCGCGCGAGCGCTGGCGCTGGTGCAG 187
Db 121 CCGGGGACCCGGAAGATCTACCGCACCTTTGGTGGTGGCCCAATGCCTAGTGTGCATGCACTGG 180
Qy 188 CGCGGGGACCCGCGCGCTTCCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGGCCTGG 247
Db 181 GGCTCACAGCGCTCACCTGCGGACCTTTCTTCCACAGAGTGTGCATCCCTGGAAGAGCTG 240
Qy 248 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307
Db 241 GTGGCCAGGGTGTGCGAGAGCTCTGCGAGCGCAACGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 308 GTGGCGCGGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
Db 301 TTTGAGCTGCTTAAAGAGGCGAGCGCGCGCTCCCATGGCTTCACTAGTAGCGTGGCT 360
Qy 368 TTGCGCTGCTGGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427
Db 361 AGCTACTTGGCCACACTGTTATTAGACCGCTGCGTGTGCTGAGTGTGATGATGCTACTG 420
Qy 428 AGCTACTTGGCCACACTGTTATTAGACCGCTGCGTGTGCTGAGTGTGATGATGCTACTG 487
Db 421 TTGAGCGGAGTGGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 488 CTGGCGCGCGTGGCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
Db 481 CTGTGTGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 548 CTGTGTGCTGCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 607
Db 541 ACCACGAGATATCTGGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 608 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 665
Db 601 AATTTCACCTTAGTCTTACACAGATCAAGAGCAGTAGTCCGCGAGGACGACCG 660
```

QY 666 -----AACGGCCTGGAACCATAGCGTCAGGAGCGGG 700

Db 661 AAACCCCTGGCCTTGCCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACCGACTACAAGT 720

QY 701 GTCCCCCTGGCCTTGCCAGCCCGGGTGCAGGAGGCGCGGGGCGCAGTGGCCAGCCGAAGT 760

Db 721 GTGCTTACACTAAGAAGGCCAGATGCTATCTGTCTCCGAGAGTGGAGGAGGCCACCCAC 780

QY 761 CTGCGGTGCCCAGAGGCCAGCGCTGGCGTGCCTTGAGCCGGAGCGGACGCCCGTT 820

Db 781 AGGCAAGTGTACCAACCCCATCAGGCAAAATCATGGTGCCAAAGTCTGCTCGGTCCCCC 840

QY 821 GGGCAGGGGTCTCTGGGCCCAACCGGCGAGGACGCGTGGACCGAGTACCGTGTCTGT 880

Db 841 GAGGTGCTTACTCAGAGAAAGATTGTCTCTTAAGGAAGGTGTGTGACCTGAGTCTC 900

QY 881 GTGGTGTCACTGCCAGACCGCGGGAAGAACCCACTCTTTGGAGGGTGGCGTCTCTGGC 940

Db 901 TCTGGG-----TCGGTGTGCTGTAACACAAAGCCAGCTCCACATCTCTGCTG 948

QY 941 ACGGGCCACTCCCAACCCATCCGTGGGCGCGCAGCACACCGGGGCCCCCATCCACATCG 1000

Db 949 TCACACCCCGCCAAATGCTTTCAGCTCAGGCCAATTTATTGAGACACAGACATTTCCCT 1008

QY 1001 CGGCCACCACTCCCTGGGACACGCTTGTCCCGGTGTACGCCGAGACCAAGCACTTC 1060

Db 1009 TACTCAGGGGAGATGCCAAGAGCGTCTAAACCCCTCATCTCTTACTCAGACACCTCCAG 1068

QY 1061 CTCTACTCTCTCAGCGGACAAGGAGCAGCTCGGCCCTCTCTTACTCAGCTCTCTGAGG 1120

Db 1069 CTTAACTTGACTGGGGCCAGGAGCTGGTGGAGATCATCTTTCTGGGCTCAAGGCCCTAGG 1128

QY 1121 CCAGCCTGACTGGCGCTCGGAGGCTCGTGAGACCACTTTCTGGGTTCAGGCCCTGG 1180

Db 1129 ACATCAGGACCACTCTCAGGACACACGCTCTATCGCGTGCATCTAGGAGATGCGGCC 1188

QY 1181 ATGCCAGGACTCCCCCAGCTTGGCCCGCTGCCCGAGCGCTACTGGCAATGCGGCC 1240

Db 1189 CTGTTCCAAAGCTGTGTGTAACCATGACAGTGCAGTGCACAAATATGTGACACTCTCAGGTCA 1248

QY 1241 CTGTTCTGAGGTGCTTTGGGAACACAGCGCAGTGCCTTACGGGGTCTCTCTCAAGACG 1300

Db 1249 CATTGAGGTTTCGACAGCAACACACAGTGCAGATGCCCT----- 1292

QY 1301 CACTGCCGCTGCGAGTGGGTACCCCGACAGCGCGGTGCTGTGCCCGGGAGAGCC 1360

Db 1293 -----GAACACCAAGCCCGCAGCCAGCCACTCATGGATTG 1323

QY 1361 CAGGGCTCTGTGGGGCCCCGAGGAGGAGACACAGACCCCGTGCCTTGGTGACGTG 1420

Db 1324 CTCGGCTGCACAGCACTCCCTGGCAGGTATATGGTTTTCTTCGGGCGCTCTCTGCAAG 1383

QY 1421 CTCGGCAGCACAGACGCCCCCTGCGAGGTACGGCTTCGTGGGGCGTGCCTGCGCGG 1480

Db 1384 GTGCTGCTCTAGTCTCTGSGGTACAGGCACAATGAGCGCGCTCTTTAAGAACTTA 1443

QY 1481 CTGGTGCCCCAGGCTCTGGGGTCCAGGCACAAGAACCGCGCTTCTCAGGAACACC 1540

Db 1444 AAGAAGTTTCACTCGTTGGGGAATACGGCAAGCTATCACTGAGGAACTGATGTGGAG 1503

QY 1541 AAGAAGTTTCACTCTGGGAGCATGCCAAGCTCTCGTGGCAGGAGCTGACGTGGAG 1600

Db 1504 ATGAAGTAGAGGATGCGCACTGGCTCCGAGCAGCGCCGGGAGGACCGTGTCCCCGCT 1563

QY 1601 ATGAGCTGCGGGACTCGCTTGGCTCGCAGGAGCCAGGGTGTGGCTGTGTTCGGCC 1660

Db 1564 GCAGACACCGCTGAGGAGAGGATCCTGGCTACGTTCTGTCTGCGGTGAGGACACA 1623

QY 1661 GCAGAGCACCGCTGCGGTGAGGAGATCCTGGCCAAGTTCTGCACTGSGCTGATGAGTGTG 1720

Db 1624 TACGTGTACAGTGTCTAGGTCTATCTTTTACATCACAGAGCACATTCAGAGAAAC 1683

QY 1721 TACGTGTACAGTGTCTAGGTCTTTCTTTTATGTACAGGAGACGATTTTCAAAAGAAC 1780

Db 1684 AGCCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCAGAGCATTTGGAGTCAAGCAA 1743

QY 1781 AGCCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGTGTGCAAGCATTTGGAATCAGACAG 1840

Db 1744 CACTTTGAGAGAGTGGCGCTACGGGAGCTGTCAACAAGAGAGTTCAGGCATCACCAGGAC 1803

QY 1841 CACTTGAAGAGGTGGAGCTGCGGGAGCTGTGCGAAGCAGAGGTTCAGGAGCATCGGGAA 1900

Db 1804 ACCTGCTAGCCATGCCCATCTGCAGACTGCGCTTATCCCCAAGCCCAACAGCCCTGCGG 1863

QY 1901 GCCAGGCCCGCCTGCTGAGCTCCAGACTCCGCTTATCCCCAAGCCTGACGGCTGCGG 1960

Db 1864 CCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTTTGGGCGAGAGAACAGGCC 1923

QY 1961 CCGATTGTGAACATGAGTACGTCGTGGGAGCCAGACGTTCCGAGAGAAAAGAGGCC 2020

Db 1924 CAGCATTTTACCCAGCGCTCTCAAGACTCTCTTCAGCATGCTCAACTATATGACGGGACAAA 1983

QY 2021 GAGCGTCTCACTCGAGGGTGAAGGACTGTTCAAGGTGCTCACTACGAGCGGGCGCG 2080

Db 1984 CATCTCACTTATGGGGTCTTCTGACTGGGTATGAATGACATCTACAGAGCTTGGCGG 2043

QY 2081 CGCCCCGGCCTCTGCGGCGCTCTGTGCTGGGCTTGACGATATCCACAGGCGCTGGCGC 2140

Db 2044 GCCTTTGCTGCTGCTGCTGCTGACCCAGACACACCCAGGATGACTTTGTTAAGGCA 2103

QY 2141 ACCTTGTGCTGCTGCTGCGGGCCAGGACCCCGCCTGAGCTGACTTTGTCAAGGTG 2200

Db 2104 GATGTACCGGGCGCTATGATGCCATCCCGCAGGGTAAGCTGGTGAGGTGTGTGCCAAT 2163

QY 2201 GATGTACGGGGCGTACGACACCATCCCGCAGGAGGCTCACGGAGTTCATCGGCAGC 2260

Db 2164 ATGATCAGGCACCTCGGAGAGCAGCTACTGTATCCGCGAGTATGACGTGGTCCGAGAGAT 2223

QY 2261 ATCATCAACCCC--AGAACAGTACTGCGTGGTATCCGCTGGTTCGAGAGGCC 2317

Db 2224 AGCCAAGGCCAAGTCCACAAAGTCTTTAGGAGACAGTACCACCCTCTCTGACCTCCAG 2283

QY 2318 GCCATGGGCACGTCGCCAAGGCTTCAAGAGCCAGCTCTCTACCTTGACAGACCTCCAG 2377

Db 2284 CCATACATGGCCAGTCTCTTAAGCATCTGCAGGATTCAGATGCCAGTGCACGTGAGGAAC 2343

QY 2378 CCGTACATGGAGAGTCTGCTGCTCACCTGCAGGAGAC-----CAGCCGCGTGGAGAT 2431

Db 2344 TCCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAGCAGCAGCGCTGTTTGTACTTC 2403

QY 2432 GCGTGTCTATCGAGCAGAGCTCTCTCCCTGAATGAGCCAGCAGTGGCTCTTCGACGTC 2491

Db 2404 TCCCTGCATCTCTGCTCACAGTGTCTGTAAGATTGGTGACAGGTGCTATACGCAAGTC 2463

QY 2492 TCCCTACGCTTCTATGTGCCACACCGCGTGCATCAGGGGCAAGTCTTACGTCCAGTGC 2551

Db 2464 CAGGGATCCCCCAGGCTCCAGCCTTCCACCCTGCTGCGAGTCTGTGTTTCGGAGAC 2523

QY 2552 CAGGGATCCCGCAGGGCTCCATCTCTCCACGCTCTCTGACGCTGTGCTACGCGGAC 2611

Db 2524 ATGGAAACAAAGCTGTTGCTGAGGTGCAGCGGATGGGTGCTTTTACGTTTGTGTGAT 2583

QY 2612 ATGGAAACAAAGCTGTTGGGGGATTCGGCGGAGCGGGCTGCTCTGCTGTTGGTGGAT 2671

Db 2584 GACTTTCTTGTGTGAGCCTCACTTTGGACCAAGCAAAACCTTCTCAGCACCCCTGGTC 2643

QY 2672 GATTTCTTGTGTGACACTCACTTCCACCCACGCGAAACCTTCTCAGAGACCCCTGGTC 2731

Db 2644 CATGGGTTCTCAGTATGGTGCAATGAACCTTCAGAGACAGTGGTGAATCCCT 2703

QY 2732 CGAGGTGCTCCCTGAGTATGCTGCTGGTGAAGTTCGGAAGACAGTGGTGAATCCCT 2791

Db 2704 GTGAGCCTGTTACCCTGGGTGGTGGAGTCCATACACGCTGCTCTCAGTCCCTGCTT 2763

QY 2792 GTAGAAGACAGGCCCTTGGGTGGCAGCGCTTTTGTTCAGATGCCCGCCACCGCCTATTTC 2851





QY 652 GCGTCTGGGATGCG-----AACGGGCTTGAACCATAG 684  
Db 674 TCGCAGGAAGCAGCGAAACCCCTTGGCCCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 733  
QY 685 CGTCAGGAGCCCGGGGTCCCTTGGCCCTCCAGCCCGGGTGGAGAGGCGCGGGG 744  
Db 734 TCTCACCAGTACAAGTGTGCTTACGCTTAAGAAGGCCAGATGCTATCTGTCTCCGAGAGT 793  
QY 745 CAGTGCCAGCGAAGTCTGCCGTGTGCCCAAGAGGCCAGGCGTGGCGTGCCTTGAGCC 804  
Db 794 GGAGGAGGACCCACAGCAGAGGTGCTTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 853  
QY 805 GGAGCGGACGCCGTGGGACAGGGGTCTCTGGGCCACCCCGGAGGACGCGTGGACCGAG 864  
Db 854 TCCGTCTCGGTCCCGGAGGTGCTTACTTGCAGAGAAAGATTGTCTTAAAGAAAGGT 913  
QY 865 TGACCGTGGTTCTGTGTGGTGTACCTTGCAGAGCCCGGCAAGAACCACTCTTTTGA 924  
Db 914 GTCTGACCTGAGTCTCTCTGGG-----TCGGTGTGCTGTAACACAAAGCCAG 961  
QY 925 GGGTGGCTCTTGGCAGCGGCCACTCCACCCATCCGTGGGCCCGCAGCACAGCGGG 984  
Db 962 CTCACATCTCTGTGTGTCACCAACCCGCCAAATGSCCTTTCAGCTCAGGCCCATTTATGA 1021  
QY 985 CCCCCATCCACATCGCGGCCACACAGCTCCCTGGGACAGCGCTTGTCCCGGCTGTAGCG 1044  
Db 1022 GACCAAGATTTCTTTACTTCAGGGGAGATGGCCAGAGGCTTAAACCCCTCATTTCT 1081  
QY 1045 CGAGACCAAGCACTTCTCTACTCTCAGGGCAAGAGGAGCTGCGGCCCTCTCTCT 1104  
Db 1082 ACTCAGCAACCTCAGCGCTAACTTACTGAGCGGCCAGAGACTGTGGAGATCATCTTCT 1141  
QY 1105 ACTCAGCTCTGAGGCCACGCTGACTTGGCGCTCGGAGGCTGTGGAGACCATCTTCT 1164  
Db 1142 GGGCTCAAGGCTTAGGACATCAGGACCACTCTGCAGGACACACCGCTCATCGCGTCGATA 1201  
QY 1165 GGGTTCAGGCCCTTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCAGCGCTA 1224  
Db 1202 CTGGCAGATGGGCCCTTTCACACAGCTGTGTGTGAACCATGCAGAGTGCCCAATATGT 1261  
QY 1225 CTGGCAATGGGCCCTTCTTCTGGAGCTGCTTGGGAACACCGCGAGTGCCTCTAGG 1284  
Db 1262 CAGACTCCTCAGGTACATTCAGGTTTCAACAGCAACCAACAGGTGACAGATGCCCTT 1321  
QY 1285 GGTGCTCTCAAGAGCACTGCCCGCTCGGAGCTGCGTCAACCCAGCAGCGGTGTG 1344  
Db 1322 -----GAACACAGCCCAACC 1336  
QY 1345 TGCCCGGAGAACCCAGGCTCTGTGGCGCCCCGAGGAGGACAGACAGCCCGC 1404  
Db 1337 GCACCTCATGATTGTCTCCGCTGCACAGCAGTCCCTTGGCAGGTATATGTTTCTTCG 1396  
QY 1405 TCGCTGTGAGCTGTCTCCGCCAGCAGCAGAGCCCTTGGCAGGTGTACGGTTCGTGCG 1464  
Db 1397 GGCCTGTCTGCAAGGTGGTGTCTGTAGTCTCTGGGGTACCAGGCAATGAGCGCG 1456  
QY 1465 GGCCTGTCTGCGCGGTGGTGTCCCGCAGGCTCTGGGGCTCCAGGCAACAGCAACGCG 1524  
Db 1457 CTCTCTTAAAGAACTTAAAGAACTTCACTCTGTTGGGAAATACGCCAAGCTATCACTCA 1516  
QY 1525 CTCTCAGGAACACCAAGATTCATCTCCTGGGGAAGCATGCCAAGCTCTCGCTGCA 1584  
Db 1517 GAACTGATGTGAAGATGAAGTATAGAGATGGCCACTTGGCTCCGACAGCCCGGGAA 1576  
QY 1585 GGAGCTGACGTGAAGATGAGCGTGGGAGTGGCGCTTGGCTGGCAGGACCCAGGSGT 1644  
Db 1577 GGACCTGTCCCGCTGCAGAGCCGCTCTGAGGAGAGATCTTGGGTACGTTCTCTGTT 1636  
QY 1645 TGGCTGTCTTCGCGCCGAGAGCAACCTCTCGGTGAGAGATCTTGGCAAGTTCCTGCA 1704  
Db 1637 CTGGCTGATGACACATACGTGGTACAGCTGCTTAGGTTCATTTTATACACACAGAG 1696  
QY 1705 CTGGCTGATGATGTGTACGTCTGCGAGCTGCTCAGGCTTCTTTTATGTCAGGAGAC 1764

Db 1697 CACATTCAGAGAAACAGGCTCTTCTTACCCTAAGAGTGTGTGGAGCAAGCTGCAAG 1756  
QY 1765 CACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCAAG 1824  
Db 1757 CATTGAGTACAGCAACACACCTTGAAGAGTGCCTACGGGAGCTGCACAAAGAGGAGT 1816  
QY 1825 CATTGGAATCAGACAGCACTTGAAGAGGTGCAGCTCGGGAGCTGTCCGAGAGAGGT 1884  
Db 1817 CAGCATCACACAGGACACCTGGCTAGCCATGCCCATCTGCAGACTGCGCTTTCATCCCAA 1876  
QY 1885 CAGCAGCATCGGACGCCAGGCGCCCTGCTGACGTCCAGACTCCGCTTTCATCCCAA 1944  
Db 1877 GCCCAAGGCGCTCGGCGCCATTTGAACATGAGTTATAGCATGGGTACCAAGAGCTTTGGG 1936  
QY 1945 GCCTGAGCGGCTCGGCGCGATTTGNAACATGGACTAGCTCGTGGGAGCCAGAGCTTCCG 2004  
Db 1937 CAGAAGAACGAGCGCCAGCATTTACCCAGCGCTCTCAAGACTCTCTTACAGCATGCTCAA 1996  
QY 2005 CAGAGAAAGAGGCGCGAGGCTCTCACCTCGAGGCTGAGGCACCTGTTACGCTGCTCAA 2064  
Db 1997 CTATGAGCGGACAAAACATCCTCACCTTATGGGTCTTCTGTACTGGGTATGAATGACAT 2056  
QY 2065 CTACAGCGGCGCGCGCCCTTCTGGCGCTCTGTGCTGGCGCTGGAGGATAT 2124  
Db 2057 CTACAGACCTTGGCGGCGCTTTGTGCTGTGCTGTGCTGTGACCAAGACACCCAGAT 2116  
QY 2125 CCACAGGCGCTGGCGCACCTTCTGCTGTGCTGTGCGGCCAGGACCCGCGCTGAGCT 2184  
Db 2117 GTACTTTGTTAAGCAGATGTACCGGGGCTATGATGCCATCCCATCCCGAGGTAGCTGGT 2176  
QY 2185 GTACTTTGTCAAGTGTGACGGCGGTACGACACCATCCCGAGGACAGGCTAC 2244  
Db 2177 GGAGGTTTGTGCCAATATGATCAGGCACTCGGAGAGCACGTACTGTATCCCGCAGTAGC 2236  
QY 2245 GGAGTCTATGCCAGCATCATCAAAACCC---AGAACACGTACTGCTGCTGCTATGC 2301  
Db 2237 AGTGGTCGGAGAGATAGCCNAGGCCAAGTCCCAAGTCTCTTTAGGAGACAGGTACCCAG 2296  
QY 2302 CGTGGTCAGAAGCCCGCCATGGGCACGTCGCGAAGSCCTTCAAGAGCCACGCTCTCTAC 2361  
Db 2297 CTTCTGACCTCCAGCATACATGGCGGCTTCTTAAAGCATCTGCAGGATTCAGATGC 2356  
QY 2362 CTTGACAGACCTCCAGCGGTACATGCGACACTTCGTGGCTCACCTGCAGGAGAC----- 2415  
Db 2357 CAGTGCACTGAGGAACCTCCGTTGTCATCGACAGAGATCTCTATGAATGAGAGCAGCAG 2416  
QY 2416 CAGCCCGCTGAGGATGCCGTGCTCATCGACAGAGCTCCTCCCTGAATGAGGCCAGCAG 2475  
Db 2417 CAGCCTGTTGACTTCTCTGCACTTCTGCTGCTCAGAGTGTCTGTAAGATTTGTTGACAG 2476  
QY 2476 TGGCCTCTTCCAGCTCTTCTTACGCTTCATGTGCCACACCGCCGTGGCATCAGGGGCAA 2535  
Db 2477 GTGCTATACAGTGCAGGCGCATCCCGGCGCTCCAGGCTATCCACCTGCTCTGCAG 2536  
QY 2536 GTCTACGTCAGTGCAGGGGATCCCGCAGGGGTCCACTCTCTCCAGCTGCTCTGCAG 2595  
Db 2537 TCTGTGTTCCGAGACATGGAGAAAGCTGTTTGTGAGGTGCAGCGGGATGGTTGCT 2596  
QY 2596 CTTGTGCTACGGCGACATGAGAAACAGCTGTTTGGCGGGATTCGGCGGAGCGGCTGCT 2655  
Db 2597 TTTACGTTTGTGATGACTTCTTGTGGTACGCCCTCAGTTGGACCAAGCAAAACCTT 2656  
QY 2656 CTTGCTTGTGATGATTTCTTGTGGTGACACCTCACCTCCACCCAGCGGAAACCTT 2715  
Db 2657 CTTACAGACCTTGTCTTCCAGGCTTCTGAGTATGGGTGCAATGAATACTTGCAGAGAC 2716  
QY 2716 CTTACAGACCTTGTCTCCGAGGCTGCTCTGAGTATGGGTGCTGAGTATGCGGAGAG 2775  
Db 2717 AGTGGTGAACCTTCCCTGTGGAGCCTGTGTACCTGGGTGGTGCAGCTCCATACCAAGTGC 2776  
QY 2776 AGTGGTGAACCTTCCCTGTAGNAGAGAGGCGCTGCGGTGGCAGCGCTTTTGTTCAGATGCC 2835

```
Db 2777 TGCTCACTGCCTGTTTCCTGCTGGTGGCTTCTGCTGGACACTCAGACTTTGGAGTCTT 2836
QY 2836 GCCCACCAGCCTATTCCCTGGTGGCGCTGCTGTGATACCCGGACCTGGAGGTGCA 2895
Db 2837 CTGTGACTACTCAGGTTATGCCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAGTGT 2896
QY 2896 GAGCAGCTACTCCAGCTATGCCCCGACCTCCATCAGAGCCAGTCTCAGCTTCAACCGGG 2955
Db 2897 CTTCAAAGCTGGGAAGACACTCGGGAACAAGCTCTCTGCGGTCTTGGGTTGAAGTGCA 2956
QY 2956 CTTCAAAGCTGGGAAGACACTCGGTCGCAAACTCTTTGGGGTCTTGGCGCTGAAGTGCA 3015
Db 2957 CGGTCTATTCTAGACTTGCAGGTCAAGAGCTCCAGACAGTCTCATCAATATATACAA 3016
QY 3016 CAGCCTGTTTCTGGATTTGCAAGTGAACAGCTCCAGACGGTGTGCAACACACTCAAA 3075
Db 3017 GATCTTCTCTGCTTCAAGGCTCAGAGTCCATGCTGATGATTCAGCTTCCCTTTTGACCA 3076
QY 3076 GATCTTCTCTGCTGAGGCTAGAGTTCACGCAATGCTGCTGCACTCCCATTTTCATCA 3135
Db 3077 GCGTGTAGGAAGACCTACATTTCTTCTGGGCAATCATCTCCAGCCCAAGCATCTGCTG 3136
QY 3136 GCAAGTTTGGGAAGACCCACATTTTCTGCGGCTCATCTCTGACAGGGCTCCCTCTG 3195
Db 3137 CTATCTATCTGAGGTCAAGTCCAGGATGACACTAAAGGCTCTGGCTCT---- 3192
QY 3196 CTATCTCATCTCTGAAGGCAAGACGAGGATGTCGTGGGGGCCAAGGGCGCGCGCG 3255
Db 3193 -----TTCTCTCTGAAGCGGCACATTTGCTCTGCTACAGGCTTCTCTCAAGCTGGC 3247
QY 3256 CCCTCTGCTCTCGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3315
Db 3248 TGCTCTATCTGTCATCTCAAAATGCTCTCTGCGGACCTCTGAGGACAGCCCAAAACTGCT 3307
QY 3316 TCGACACCGTGTCACTAGTCCACTCTGCGGCTCACTCAGGACAGCCACAGCAGCT 3375
Db 3308 GTCGCGGAGCTCCAGAGGCGACATGACCATCTTAAAGTGCAGTGCAGCCAGCCCT 3367
QY 3376 GAGTGGAGCTCCCGGAGACAGCTGACTGCTGCTGAGGCGCGCAACCCCGGCACT 3435
Db 3368 AAGCACAGACTTTCAGACACTTTTGACT 3396
QY 3436 GCCTCAGACTTCAGACCATCTTGACT 3464

RESULT 6
LOCUS AF029235 201 bp mRNA ROD 12-NOV-1997
DEFINITION Mus musculus telomerase catalytic subunit mRNA, partial cds.
ACCESSION AF029235
NID 92605902
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.I.
TITLE Partial sequence of Mus musculus telomerase catalytic subunit
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201)
AUTHORS Drissi,R. and Cleveland,J.I.
TITLE Direct submission
JOURNAL Submitted (07-OCT-1997) Biochemistry, St Jude Children's Research
Hospital, 332 North Lauderdale, Memphis, TN 38105, USA
FEATURES
source 1..201
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="32D.3"
<1..>201
/codon_start=1
/product="telomerase catalytic subunit"

CDS
```

```
/db_xref="PID:g2605903"
/translation="FEIVTSEFQKNRLFFYKSVWSKLSQISGVROHLERVRLRELSQ
EEVRRHODTWLAMPICRLRIFP"
BASE COUNT 51 a 54 c 55 g 41 t
ORIGIN
Query Match 3.4%; Score 129; DB 28; Length 201;
Best Local Similarity 82.1%; Pred. No. 7.48e-73;
Matches 165; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
Db 1 TTCTTTTATGTCACGGAGAGACATTCAGAGAACAAGGCTCTTCTTCTACCGTAAGAGT 60
QY 1745 TTCTTTTATGTCACGGAGAGACATTCAGAGAACAAGGCTCTTCTTCTACCGTAAGAGT 1804
Db 61 GTGTGGAGCAAGCTGTCAGAGCATTTGGAGTCAGGCAACACCTTGAGAGAGTTCGGCTACGG 120
QY 1805 GTCTGGAGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGCAGCTCGG 1864
Db 121 GAGCTGTCAAGAGAGGTCAGGCATCAGCAGACACCTGGCTAGCATGCCATCTGC 180
QY 1865 GAGCTGTGGAAGCAGAGGTCAGGCATCAGCAGCACTGGGAAGCCAGGCCCTCTGACGTCC 1924
Db 181 AGACTCCGCTTCATCCCAAG 201
QY 1925 AGACTCCGCTTCATCCCAAG 1945

RESULT 7
LOCUS I66494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION I66494
NID 92724471
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 1.8%; Score 67; DB 21; Length 7218;
Best Local Similarity 2.4%; Pred. No. 5.33e-26;
Matches 9; Conservative 211; Mismatches 153; Indels 0; Gaps 0;
Db 1055 GGAGCTTGGCATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1114
QY 922 GGAGGTTGGCTCTCTGGCAGCGCCACTCCACCCATCGTGGCGCCGACGACGCG 981
Db 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
QY 982 GGGCCCCCATCCACATCGGCGCCACGCTCCCTGGGACACGCTGTCTCCCGGTGA 1041
Db 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
QY 1042 CGCGAGACCAAGCACTTCTCTACTCTCAGCGCAAGAGCAGCTCGGCGCCCTCT 1101
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 1102 CCTACTCAGCTCTCTGAGGCCAGCCTGACGTGCGCTCGGAGGCTCGTGGAGACCATCTT 1161
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 1162 TCTGGTTCAGGCGCTGATGCCAGGACTCCCGGAGGTTGCCCGCTCGCCGACG 1221
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
QY 1222 CTACTGGCAATGGCGCCCTCTTCTGGAGCTCTTGGGAACCAAGCGGAGTGCCTCA 1281
```



Db 1415 YYYYYYYYYYYY 1427  
Qy 1282 CGGGTGCTCCTC 1294

RESULT 8 I66494 7218 bp DNA PAT 23-DEC-1997  
LOCUS Sequence 14 from patent US 5670367.  
DEFINITION I66494  
ACCESSION g2724471  
NID  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7218)  
AUTHORS Dörner, F., Schelflinger, F. and Falkner, F. Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES Location/Qualifiers

source 1. .7218  
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others  
ORIGIN

Query Match 1.3%; Score 49; DB 21; Length 7218;  
Best Local Similarity 2.0%; Pred. NO. 1.39e-13;  
Matches 6; Conservative 165; Mismatches 122; Indels 0; Gaps 0;

Db 1150 YYYYYYYYYYYY 1209  
Cp 2080 CCGCGCGGCTGAGTTGACGACGCTGACGAGCTCCACTCGAGTGACGCTC 2021  
Db 1210 YYYYYYYYYYYY 1269  
Cp 2020 GGCCCTCTTCTCTCGGACGCTCTGGCTCCACGACGAGTCCATGTTCAACATCGG 1961  
Db 1270 YYYYYYYYYYYY 1329  
Cp 1960 CCGACGCCGTCAGGCTGGGATGAAGCGGAGTCTGACGCTCAGCAGGGCGGCTGGC 1901  
Db 1330 YYYYYYYYYYYY 1399  
Cp 1900 TTCCGATGTCGCTGACCTGCTCCGACGCTCCGACGCTCCGACGCTCTTCAAGTG 1841  
Db 1390 YYYYYYYYYYYY 1442  
Cp 1840 CTGCTGATTCATGCTTTCCAACTTCTCCACACTCTCTCCGGTAGAAA 1788

RESULT 9  
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998  
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.  
ACCESSION AC005369  
NID 93367505  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 74371)  
AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,  
Rojeski, H., Miguel, T., Miller, C., Pittluck, S., Pollard, M.,  
Subramanian, S. and Martin, C.H.

TITLE Sequencing of human chromosome 5  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 74371)  
AUTHORS Ricke, D.O.  
TITLE Large Scale Sequence Analysis and Annotation with the Sequence  
Comparison Analysis (SCAN) System  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 74371)

AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,  
Davis, C.A., Kadner, K., Miguel, T., Pittluck, S., Pollard, M.,  
Rojeski, H., Subramanian, S. and Martin, C.H.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome  
Institute, Lawrence Berkeley National Laboratory, MS 74-157,  
Berkeley, CA 94720, U.S.A.  
COMMENT Sequence submitted by:  
DOE Joint Genome Institute.  
FEATURES Location/Qualifiers  
source 1. .74371  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="5q"  
/clone="119j3"  
/chromosome="5"  
/note="LBNL H175"  
893. .1030  
/rpt\_family="Alu"  
repeat\_region 2295. .2438  
/rpt\_family="Alu"  
repeat\_region 2818. .2859  
/note="(GT)21"  
/rpt\_type=tandem  
/rpt\_unit=GT  
misc\_feature join(3246. .3410,3721. .3828)  
/standard\_name="RLF"  
/note="65% & 69% protein identity GenPept:U02377"  
3431. .3724  
/rpt\_family="Alu"  
repeat\_region 3707. .3728  
/note="(A)22"  
/rpt\_type=tandem  
/rpt\_unit=A  
repeat\_region 4366. .4661  
/rpt\_family="Alu"  
repeat\_region 5327. .5602  
/rpt\_family="Alu"  
repeat\_region 6586. .6956  
/rpt\_family="L1"  
repeat\_region 6647. .6684  
/note="(CA)19"  
/rpt\_type=tandem  
/rpt\_unit=CA  
repeat\_region 7113. .7373  
/rpt\_family="Alu"  
complement(7830. .8185)  
/standard\_name="possible repeat"  
8258. .8503  
/rpt\_family="Alu"  
repeat\_region 9070. .9387  
/rpt\_family="Alu"  
repeat\_region complement(9740. .9845)  
/rpt\_family="MER42"  
repeat\_region complement(10440. .11015)  
/rpt\_family="Alu"  
repeat\_region 11950. .12250  
/rpt\_family="Alu"  
repeat\_region 12057. .12085  
/note="(A)29"  
/rpt\_type=tandem  
/rpt\_unit=A  
repeat\_region 12365. .12645  
/rpt\_family="Alu"  
repeat\_region 13727. .13750  
/note="(AC)12"  
/rpt\_type=tandem  
/rpt\_unit=AC  
repeat\_region 13783. .14024  
/rpt\_family="L1"  
repeat\_region 14175. .14470  
/rpt\_family="Alu"  
complement(14906. .15259)  
repeat\_region



QY 131 CTGCGCTGCCACGTCGTCGGCGCGCTGGCGCCCGCGCTGGCGCTGGCGCGC 190  
Db 143 SRNMGASVGSDRNTRKHAKNSADGVSKNGNDRNRYGTGTGKSNVSNNGGGRDVS 202  
QY 191 GGGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTGCCTGTGTGCGCGCCCTGGGAC 250  
Db 203 SYANNKCGSSC 214  
QY 251 GCACGCGCGCC 262

RESULT 11  
LOCUS 128278 215 bp DNA PAT 30-OCT-1996  
DEFINITION Sequence 5 from patent US 5569830.  
ACCESSION 128278  
NID g1819054  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 215)  
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.  
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease  
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;  
FEATURES Location/Qualifiers  
source 1..215  
BASE COUNT 15 a 8 c 25 g 26 t 141 others  
ORIGIN  
Query Match 0.98; Score 35; DB 21; Length 215;  
Best Local Similarity 13.28; Pred. No. 8.17e-05;  
Matches 26; Conservative 76; Mismatches 95; Indels 0; Gaps 0;

Db 6 MSSSVVSRASCNDXAKKDGNTTSSWTTCCNRTGVCVCTDTTYRVNNDGHNKYSAN 65  
Cp 1616 AGTCCCGCAGCTCATCTCCACGTACGTCCTCGACGAGAGCTTGGCATGCTCCCCA 1557  
Db 66 YNTGGNVGAATHYTHTVNSGADSTVTDVSYNASGTSSNGTGNRSGADSYSSKT 125  
Cp 1556 GGAGATGAACTTCTGCTTCTGAGGAGCGGCTTCTGTGCTGCGGACCCCGAGA 1497  
Db 126 AMFSRNRTKTANNVDSVGSDRNTRKHAKNSADGVSKNGNDRNRYGTGT 185  
Cp 1496 GGCTGGGGCAGCAGCGCGCGCAGGCGCCGACGAGCGGTACACCTGCGAGGGC 1437  
Db 186 KSNVSNNGGGRDVS 202  
Cp 1436 TGCTGTGCGCGAGC 1420

RESULT 12  
LOCUS AC005369 74371 bp DNA PRI 01-AUG-1998  
DEFINITION Homo sapiens chromosome 5, BAC clone 119j3 (LBNL H175), complete sequence.  
ACCESSION AC005369  
NID g3367505  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 74371)  
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.  
TITLE Sequencing of human chromosome 5  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 74371)  
AUTHORS Ricke,D.O.  
TITLE Large Scale Sequence Analysis and Annotation with the Sequence

Comparison Analysis (SCAN) System  
Unpublished  
3 (bases 1 to 74371)  
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.  
COMMENT Sequence submitted by:  
DOE Joint Genome Institute.  
FEATURES Location/Qualifiers  
source 1..74371  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="5q"  
/clone="119j3"  
/chromosome="5"  
/note="LBNL H175"  
893..1030  
/rpt\_family="Alu"  
2295..2438  
/rpt\_family="Alu"  
2818..2859  
/note="(GT)21"  
/rpt\_type=tandem  
/rpt\_unit=GT  
join(3246..3410,3721..3828)  
/standard\_name="RLF"  
/note="65% & 69% protein identity GenPept:U22377"  
3431..3724  
/rpt\_family="Alu"  
3707..3728  
/note="(A)22"  
/rpt\_type=tandem  
/rpt\_unit=A  
4366..4661  
/rpt\_family="Alu"  
5327..5602  
/rpt\_family="Alu"  
6586..6956  
/rpt\_family="L1"  
6647..6684  
/note="(CA)19"  
/rpt\_type=tandem  
/rpt\_unit=CA  
7113..7373  
/rpt\_family="Alu"  
complement(7830..8185)  
/standard\_name="possible repeat"  
8258..8503  
/rpt\_family="Alu"  
9070..9387  
/rpt\_family="Alu"  
complement(9740..9845)  
/rpt\_family="MER42"  
complement(10440..11015)  
/rpt\_family="Alu"  
11950..12250  
/rpt\_family="Alu"  
/rpt\_type=tandem  
12057..12085  
/note="(A)29"  
/rpt\_type=tandem  
/rpt\_unit=A  
12365..12645  
/rpt\_family="Alu"  
13727..13750  
/note="(AC)12"  
/rpt\_type=tandem  
/rpt\_unit=AC  
13783..14024  
/rpt\_family="L1"



```

BASE COUNT      16 a      66 c      51 g      22 t      46 others
ORIGIN
LTAQAVRRAAREITRIPLYKKG"

Query Match      0.8%; Score 30; DB 21; Length 201;
Best Local Similarity 36.1%; Pred. No. 5.30e-02;
Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

Db 2 TGCCSACTGSCGSGTSCSCGSGAGTCSCCGSCCSCAGCCGCGGCGCGCCGSCGSC 61
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 27 TGCTGGCCACGTGGAGAGCCTGCGCCCGCCACCCCGCGGATGCGCGCGCTCCCGCT 86
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CSCTGTCGTCGSCGSCGTCGTCGAGGCGGCGSCGTCGTCGTCGTCGTCGTCGTC 121
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 GCGAGCGCTGCGCTGCTGCTGCGAGCCACATACCGGAGGTGCTGCGCTGCGCCAGT 146
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 CSGCTCSCGCTSCAGGCGCCAGG 145
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 TCCTGGCGCGCTGGGCGCCAGG 170

RESULT 14
LOCUS      A10158      201 bp      DNA      PAT      01-DEC-1993
DEFINITION S.griseus gene for preprotease (partial).
ACCESSION  A10158
NID         9490196
KEYWORDS   Streptomyces griseus.
SOURCE     Streptomyces griseus
ORGANISM   Streptomyces griseus
            Eubacteria; Firmicutes; Actinomycetes; Streptomycetes;
            Streptomycetaceae; Streptomycetes.
REFERENCE  Garvin,R.T. and James,E.
            Production of active proteins containing cystine residues
            Patent: EP 0222279-A 2 20-MAY-1987;
            Cangene Corporation
FEATURES   Location/Qualifiers
            source      1..201
                        /organism="Streptomyces griseus"
                        /db_xref="taxon:1911"
            CDS         1..>201
                        /codon_start=1
                        /transl_table=11
                        /product="preprotease"
                        /db_xref="PID:g490197"
                        /translation="MPHSPVSPAESPAPQPRPVVSRRLLEGGAAVLGALASAP
LTAQAVRRAAADEPWNDFG"
BASE COUNT      13 a      66 c      54 g      22 t      46 others
ORIGIN

Query Match      0.8%; Score 30; DB 21; Length 201;
Best Local Similarity 36.1%; Pred. No. 5.30e-02;
Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

Db 2 TGCCSACTGSCGSGTSCSCGSGAGTCSCCGSCCSCAGCCGCGGCGCGCCGSCGSC 61
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 27 TGCTGGCCACGTGGAGAGCCTGCGCCCGCCACCCCGCGGATGCGCGCGCTCCCGCT 86
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CSCTGTCGTCGSCGSCGTCGTCGAGGCGGCGSCGTCGTCGTCGTCGTCGTCGTC 121
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 GCGAGCGCTGCGCTGCTGCTGCGAGCCACATACCGGAGGTGCTGCGCTGCGCCAGT 146
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 CSGCTCSCGCTSCAGGCGCCAGG 145
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 TCCTGGCGCGCTGGGCGCCAGG 170

RESULT 15
LOCUS      A10162      201 bp      DNA      PAT      01-DEC-1993
DEFINITION Synthetic DNA for preprotease leader & prochymosin.
ACCESSION  A10162
NID         9489115
KEYWORDS

```

```

SOURCE     synthetic construct.
ORGANISM   synthetic construct
            artificial sequence.
REFERENCE  1 (bases 1 to 201)
AUTHORS   Garvin,R.T. and James,E.
TITLE     Production of active proteins containing cystine residues
JOURNAL   Patent: EP 0222279-A 6 20-MAY-1987;
FEATURES   Location/Qualifiers
            source      1..201
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
BASE COUNT      22 a      51 c      66 g      16 t      46 others
ORIGIN

Query Match      0.8%; Score 30; DB 21; Length 201;
Best Local Similarity 36.1%; Pred. No. 5.30e-02;
Matches 52; Conservative 35; Mismatches 57; Indels 0; Gaps 0;

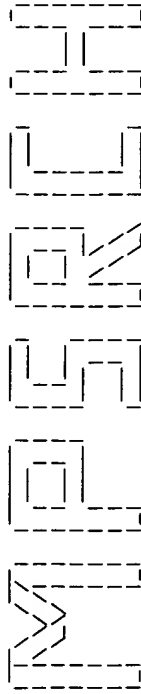
Db 57 CTTGGCCCTTSGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 116
  ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 170 CTTGGGCCCCAGGCGCGCAGCAACGTGCGCAGCGGAGGAGGAGGAGGAGGAGG 111
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 CSAGSAGSCGSCGSGASACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 176
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 110 GCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 51
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 CSGGSGASACSGGSGAGTGSGGCA 200
  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 50 CCAGGCTTCCACACGTGCGCAGCA 27

Search completed: Fri Dec 25 01:59:16 1998
Job time : 23728 secs.

```

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 01:59:35 1998; MasPar time 10944.33 Seconds  
Tabular output not generated. 621.452 Million cell updates/sec

Title: >US-08-951-733-19  
Description: (1-3798) from US08951733.seq  
Perfect Score: 3798  
N.A. Sequence: 1 CCACGGGTCCGGCAGCGCT.....GGAATAGTCATCCCTGAT 3798  
Comp: GGTGCGCAGCGCGTCGCGA.....CCTATCAGGTAGGCGACTA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Mismatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est55  
Database: genbank-est107  
5:gb\_est1 6:gb\_est10 7:gb\_est11 8:gb\_est12 9:gb\_est13  
10:gb\_est14 11:gb\_est15 12:gb\_est16 13:gb\_est17  
14:gb\_est18 15:gb\_est19 16:gb\_est20 17:gb\_est21  
18:gb\_est21 19:gb\_est23 20:gb\_est4 21:gb\_est5 22:gb\_est6  
23:gb\_est7 24:gb\_est8 25:gb\_est9 26:gb\_gss1 27:gb\_gss2  
28:gb\_gss3 29:gb\_gss4

Statistics: Mean 12.047; Variance 2.654; scale 4.539

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	385	10.1	389	8	AA281296 EST08902.r1 NCI_CGAP_GC	0.00e+00
2	248	6.5	409	24	AA311750 EST182469 Jurkat T-cel	0.00e+00
3	82	2.2	315	13	AY02e05.s1 NCI_CGAP_GC	1.90e-88
4	84	2.2	340	13	AA811084 OA85C05.s1 NCI_CGAP_GC	7.88e-92
5	58	1.5	252	12	AA754459 97SN1787 Rice Immature	3.51e-49
6	58	1.5	252	12	AA754459 97SN1787 Rice Immature	3.51e-49
7	52	1.4	303	24	AA299878 EST12462 Uterus tumor	6.91e-40
8	49	1.3	247	12	AA754458 97SN1784 Rice Immature	2.37e-35
9	47	1.2	247	12	AA754458 97SN1784 Rice Immature	2.25e-32
10	33	0.9	375	23	AA200728 mul3h09.r1 Soares 2NbM	5.88e-13
11	33	0.9	2275	11	AF034173 Homo sapiens ntcon2 co	5.88e-13
12	31	0.8	660	11	AF034177 Homo sapiens ntcon6 co	1.86e-10

C	13	25	0.7	143	18	AI171394	EST217350 Normalized r	1.41e-03
	14	25	0.7	181	6	AA386387	EST81369 Prostate glan	1.41e-03
	15	25	0.7	187	13	AA855630	VW68h11.r1 Stratagene	1.41e-03
	16	25	0.7	213	16	R72797	Y109c09.r1 Homo sapien	1.41e-03
C	17	25	0.7	242	15	AI012107	EST206558 Normalized r	1.41e-03
	18	25	0.7	278	13	AA882818	TENS0393 T. cruzi epim	1.41e-03
	19	25	0.7	301	14	AA848961	EST191723 Normalized r	1.41e-03
C	20	25	0.7	302	15	AI010844	EST205295 Normalized r	1.41e-03
	21	27	0.7	317	27	AQ008007	CIT-HSP-228C3.TF CIT-	9.41e-06
	22	25	0.7	318	14	AA799774	EST189271 Normalized r	1.41e-03
	23	25	0.7	318	16	H50134	VO27a07.r1 Homo sapien	1.41e-03
	24	25	0.7	343	22	AA098755	T3985 MVAT4 bloodstrea	1.19e-04
	25	25	0.7	359	24	AA303595	EST16305 Aorta endothe	1.41e-03
	26	25	0.7	382	5	R54656	Y74d04.r1 Homo sapien	1.41e-03
	27	26	0.7	396	19	T26788	T529 Trypanosoma bruce	1.19e-04
	28	26	0.7	400	16	H4707	Y24c08.r1 Homo sapien	1.19e-04
C	29	25	0.7	420	17	AI071447	UI-R-C1-Ku-c-05-0-UI.s	1.41e-03
	30	25	0.7	432	6	AA043978	Z58f01.s1 Soares preg	1.41e-03
C	31	25	0.7	438	14	AA94206	UI-R-Al-dx-a-11-0-UI.s	1.41e-03
	32	25	0.7	478	14	AA925372	UI-R-Al-ee-d-07-0-UI.s	1.41e-03
C	33	25	0.7	487	16	R86860	Ym86a03.r1 Homo sapien	1.41e-03
	34	25	0.7	504	15	AI010120	EST204571 Normalized r	1.41e-03
C	35	25	0.7	511	17	AI072026	UI-R-C2-nd-e-12-0-UI.s	1.41e-03
	36	25	0.7	511	15	AA997735	UI-R-C0-hn-b-08-0-UI.s	1.41e-03
C	37	25	0.7	540	14	AA901048	UI-R-Al-dv-e-09-0-UI.s	1.41e-03
	38	27	0.7	560	12	AA775218	ac79e03.s1 Stratagene	9.41e-06
	39	25	0.7	590	10	AA622227	nc43c05.s1 NCI_CGAP.Pr	1.41e-03
C	40	27	0.7	660	11	AF034177	Homo sapiens ntcon6 co	9.41e-06
	41	25	0.7	677	18	AI105428	EST214717 Normalized r	1.41e-03
C	42	25	0.7	705	14	AA941159	L25143.Sprime LD dros	1.41e-03
	43	28	0.7	1287	12	AF038250	Homo sapiens clone ntc	6.90e-07
C	44	28	0.7	1287	12	AF038250	Homo sapiens clone ntc	6.90e-07
C	45	26	0.7	2275	11	AF034173	Homo sapiens ntcon2 co	1.19e-04

ALIGNMENTS

RESULT	1	AA281296	389 bp	mrna	EST	14-AUG-1997
LOCUS		2108902.r1 NCI_CGAP_GCB1	Homo sapiens	cdna	clone IMAGE:712562 5',	
DEFINITION		2108902.r1 NCI_CGAP_GCB1	Homo sapiens	cdna	clone IMAGE:712562 5',	
ACCESSION		AA281296	mrna	sequence.		
NID		91924194				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;				
		Homo.				
REFERENCE		1 (bases 1 to 389)				
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
		Tumor Gene Index				
JOURNAL		Unpublished (1997)				
COMMENT		Contact: Robert Strausberg, Ph.D.				
		Tel: (301) 496-1550				
		Email: Robert.Strausberg@nih.gov				
		This clone is available royalty-free through LNL; contact the				
		IMAGE Consortium (info@image.llnl.gov) for further information.				
		Insert Length: 2187 Std Error: 0.00				
		Seq primer: -28m13 rev2 Et from Amersham				
		High quality sequence stop: 385.				
		Location/Qualifiers				
		1. .389				
		/organism="Homo sapiens"				
		/note="Vector: pT73D-Pac (Pharmacia) with a modified				
		polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA				
		was prepared from human tonsillar cells enriched for				
		germinal center B cells by flow sorting (CD20+, IgD-),				
		provided by Dr. Louis M. Staudt (NCI), Dr. David Allman				
		(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was				
		primed with a Not I - oligo(dT) primer				

FEATURES

source

5'-TGTACCAATCTCAAGTGGGAGCGCGGCTCAATTTTCTTTTCTTTT-  
3'). Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9608"  
/clone="IMAGE:71562"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
<1. ->389  
87 a 102 c 123 g 77 t

BASE COUNT  
ORIGIN

Query Match 10.1%; Score 385; DB 8; Length 389;  
Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GCCAAGTTCCTGCTGCTGATGAGTGTGACGTGCTGAGCTGCTCAGTCTTTCTTT 60  
QY 1691 GCCAAGTTCCTGCTGCTGATGAGTGTGACGTGCTGAGCTGCTCAGTCTTTCTTT 1750

Db 61 TATGTCAGGAGACACCTTCAAGACAGGCTCTTTTCTACCGGAAGTGTCTGG 120  
QY 1751 TATGTCAGGAGACACCTTCAAGACAGGCTCTTTTCTACCGGAAGTGTCTGG 1810

Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGCAGCTGCGGAGCTG 180  
QY 1811 AGCAAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGCAGCTGCGGAGCTG 1870

Db 181 TCGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGCCGCCCTGCTGACGTCCAGACTC 240  
QY 1871 TCGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGCCGCCCTGCTGACGTCCAGACTC 1930

Db 241 CGCTTCATCCCAAGCTGACGGGCTGCGCGGATTTGAACATGATAGTCTGTGGGA 300  
QY 1931 CGCTTCATCCCAAGCTGACGGGCTGCGCGGATTTGAACATGATAGTCTGTGGGA 1990

Db 301 GCCAAGCTTCCGAGAAAGAGGCGGAGCTCTCACTCCAGGGTGAAGGCACTG 360  
QY 1991 GCCAAGCTTCCGAGAAAGAGGCGGAGCTCTCACTCCAGGGTGAAGGCACTG 2050

Db 361 TTCAGCGTGTCACTACGAGGGGCGCG 389  
QY 2051 TTCAGCGTGTCACTACGAGGGGCGCG 2079

RESULT 2 A311750 409 bp mRNA EST 19-APR-1997  
LOCUS EST182469 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA  
DEFINITION sequence.  
ACCESSION A311750  
NID g1964077  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;  
Homo.

REFERENCE 1 (bases 1 to 409)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,  
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,  
Glodek,A., Gnehm,C.B., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,  
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dimke,D., Feng,D.-F., Ferrrie,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,  
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,  
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,  
Fraser,C.M. and Venter,J.C.  
Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl.), 3-174 (1995)  
96026280

CONTACT: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1..409  
/organism="Homo sapiens"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"  
/db\_xref="ATCC (inhost):158964"  
/db\_xref="taxon:9606"  
/clone\_lib="Jurkat T-cells VI"  
/cell\_type="T-lymphocyte"  
<1. ->409  
65 a 120 c 133 g 86 t 5 others

BASE COUNT  
ORIGIN

Query Match 6.5%; Score 248; DB 24; Length 409;  
Best Local Similarity 99.2%; Pred. No. 0.00e+00;  
Matches 248; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 GTTTGGTGGATGATTCTTGTGTGACACTCACCTCACCCAGCGAAACCTTCCTCA 60  
QY 2661 GTTTGGTGGATGATTCTTGTGTGACACTCACCTCACCCAGCGAAACCTTCCTCA 2720

Db 61 GGACCTTGGTCCGAGGTGTCCTGAGTATGGTGGTGAACCTTCGGAAGACATGG 120  
QY 2721 GGACCTTGGTCCGAGGTGTCCTGAGTATGGTGGTGAACCTTCGGAAGACATGG 2780

Db 121 TGAATCTCCCTGTAGAGACGAGCCCTGGTGCACGGCTTTTTCAGATCCGCGCC 180  
QY 2781 TGAATCTCCCTGTAGAGACGAGCCCTGGTGCACGGCTTTTTCAGATCCGCGCC 2840

Db 181 ACGCNTATTCCCTGGTGGGCTGCTGCTGGATACCCGACCTGGAGTGCAGAGCG 240  
QY 2841 ACGCNTATTCCCTGGTGGGCTGCTGCTGGATACCCGACCTGGAGTGCAGAGCG 2900

Db 241 ACTACTCCAG 250  
QY 2901 ACTACTCCAG 2910

RESULT 3 A4748707 315 bp mRNA EST 18-FEB-1998  
LOCUS ny02e05.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1270592,  
DEFINITION mRNA sequence.  
ACCESSION A4748707  
NID g2788665  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 315)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),





Department of Cytogenetics  
National Inst. of Agri. Sci. and Tech, RDA  
Suwon, Kyunggido, Korea  
Tel: 82 331 290 0301  
Fax: 82 331 290 0307  
Email: myeun@sun20.asti.re.kr  
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji  
University, Yongin, Korea. 449-728 bnhnm@bioserver.myongji.ac.kr  
Seq primer: M13 Reverse Primer.  
Location/Qualifiers

## FEATURES

source

1. .252  
/organism="Oryza sativa"  
/cultivar="Milyang23"  
/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:  
XhoI; Directional cDNA library inserted into lambda ZAPII  
vector at 5' end with EcoRI and 3' end with xho I site."  
/db\_xref="taxon:4530"  
/clone="97SN1787"  
/tissue\_type="Immature Seed"  
/dev\_stage="5 days after pollination"  
/lab\_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

## ORIGIN

Query Match 1.5%; Score 58; DB 12; Length 252;  
Best Local Similarity 10.0%; Pred. No. 3.51e-49;  
Matches 25; Conservative 120; Mismatches 105; Indels 0; Gaps 0;

Db 1 HWDCTMTVRCGCCBAMMTTSYBCHGNBVWCVASHGNYMSVHNCTBRGTHCDCKKNVNW 60  
QY 3032 TTGAGGTGAACAGCTCCACAGCGGTGGCACCAATCTACAGATCCTCTGTCGAG 3091  
Db 61 STMTGTYNNBNSVDHMYBNVNTKVDGNHTRCSRWRBYTRMAHYHDYTNCBYNNND 120  
QY 3092 CGGTACAGGTTTCAGCGATGTGTCGAGCTCCATTCATCAGCAAGTTTGAAGAAC 3151  
Db 121 YMHMBBYBTGCTCMWCBYHNTKCTASGHNHSTNYDVKSTNTWGTBVSYSKSMH 180  
QY 3152 CCCACATTTTCTCGCGGTCTATCTCGACAGCGGCTCCCTCTGCTACTCCATCTGAAA 3211  
Db 181 GYWCBBYKHTKYSTTRATSYTCVRKYCYVMWTKKVVKKYHVVBGCHETDSKCKTMW 240  
QY 3212 GCCAAGACGAGGAGTCTGCTGGGCCCAAGGCGCGCCCTCTGCCCTCCGAG 3271  
Db 241 MTNRHVMFTST 250  
QY 3272 GCGGTGCAGT 3281

RESULT 6  
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998  
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa  
ACCESSION AA754459  
NID 92801165  
KEYWORDS EST.  
SOURCE rice.  
ORGANISM Oryza sativa  
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;  
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;  
Poales; Poaceae; Oryza.  
1 (bases 1 to 252)  
Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,  
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,  
Lee,M.C. and Eun,M.Y.  
Large-scale Sequencing Analysis of ESTs from Rice Immature Seed  
Unpublished (1998)

## REFERENCE

## AUTHORS

Contact: Eun M.Y.

Department of Cytogenetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyunggido, Korea

## TITLE

## JOURNAL

## MEDLINE

## COMMENT

Tel: 82 331 290 0301  
Fax: 82 331 290 0307

Email: myeun@sun20.asti.re.kr  
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji  
University, Yongin, Korea. 449-728 bnhnm@bioserver.myongji.ac.kr  
Seq primer: M13 Reverse Primer.  
Location/Qualifiers

## FEATURES

source

1. .252  
/organism="Oryza sativa"  
/cultivar="Milyang23"  
/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:  
XhoI; Directional cDNA library inserted into lambda ZAPII  
vector at 5' end with EcoRI and 3' end with xho I site."  
/db\_xref="taxon:4530"  
/clone="97SN1787"  
/tissue\_type="Immature Seed"  
/dev\_stage="5 days after pollination"  
/lab\_host="E. coli SOLR"

BASE COUNT 5 a 21 c 12 g 35 t 179 others

## ORIGIN

Query Match 1.5%; Score 58; DB 12; Length 252;  
Best Local Similarity 12.8%; Pred. No. 3.51e-49;  
Matches 23; Conservative 88; Mismatches 69; Indels 0; Gaps 0;

Db 24 BCHGNBYVWCVASHGNYMSVHNCTBRGTHCDCKNVNMTWGTBVSYSKSMH 83  
Cp 894 CAGGTGACACACACAGACAGGTCACCTCGCTCCACGCGCTCGCTCGCGGCGGCC 835  
Db 84 NTKYDVGNHTRCSRWRBYTRMAHYHDYTNCBYNNNDYHMHBBYBTTGCTCTMWCWB 143  
Cp 834 CAGGACCCCTGCCACAGCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 775  
Db 144 HYNKCTASGHNHSTNYDVKSTNTWGTBVSYSKSMHGYWCSBBYKVTYKVTTRATSY 203  
Cp 774 TTGGCAACGCGACACTTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 715

## RESULT

## LOCUS

AA299878 303 bp mRNA EST 18-APR-1997

EST12462 Uterus tumor I Homo sapiens cDNA 5' end, mRNA sequence.

ACCESSION AA299878

NID 91952209

EST.

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;

Homo.

REFERENCE 1 (bases 1 to 303)

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,

Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,

White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,

Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,

Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,

Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,

Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,

Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligriano,S.M.,

Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,

Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,

Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,

Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,

Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,

Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,

Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,

Fraser,C.M. and Venter,J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

COMMENT

Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.

#### FEATURES

Location/Qualifiers  
1..303

/organism="Homo sapiens"

/note="Organ: uterus; Vector: pBluescript SK-; Site\_1:

EcoRI; Site\_2: XhoI"

/db\_xref="ATCC (inhost):193066"

/db\_xref="taxon:9606"

/clone\_lib="Uterus tumor I"

/dev\_stage="adult"

<1..303

BASE COUNT 54 a 85 c 93 g 69 t 2 others

Query Match 1.4%; Score 52; DB 24; Length 303;

Best Local Similarity 98.1%; Pred. No. 6.91e-40;

Matches 53; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 CTGACGGGCTCCTCTGCTACTCTCCATCTCGAAGCCAGACGAGGTATGT 54  
|||||  
QY 3177 CTGACGGGCTCCTCTGCTACTCTCCATCTCGAAGCCAGACGAGGTATGT 3230

#### RESULT

LOCUS AA754458 247 bp mRNA EST 20-JAN-1998  
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa  
CDNA clone 97SN1784, mRNA sequence.

#### ACCESSION

AA754458

#### NID

92801164

#### KEYWORDS

EST.

#### SOURCE

rice.

#### ORGANISM

Oryza sativa

Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;

Polales; Poaceae; Oryza.

1 (bases 1 to 247)

Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

Contact: Eun M.Y.

Department of Cytogenetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyunggido, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeun@sun20.asti.re.kr

Submitted by Baek Hie Nam, Dept of Biological Science, Myongji

University, Yongin, Korea. 449-728 bhnam@bioserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

Location/Qualifiers

1..247

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db\_xref="taxon:4530"

/clone="97SN1784"

/clone\_lib="Rice Immature Seed Lambda ZAPII cDNA Library"

/tissue\_type="Immature Seed"

BASE COUNT 54 a 85 c 93 g 69 t 2 others

Query Match 1.4%; Score 52; DB 24; Length 303;

Best Local Similarity 98.1%; Pred. No. 6.91e-40;

Matches 53; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/dev\_stage="5 days after pollination"

/lab\_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others

#### ORIGIN

Query Match 1.3%; Score 49; DB 12; Length 247;

Best Local Similarity 13.0%; Pred. No. 2.37e-35;

Matches 27; Conservative 93; Mismatches 88; Indels 0; Gaps 0;

Db 28 TBWCCVRRYGTITNRKNGKRRTTTNDSCDNDNAHCRITYVWYARSKYGYGTBYISNVND 87

|||||

QY 183 TGCAGCGGGGACCCGGCGCTTTCGCGCGCTGTGGCCAGTGCCTGGTGGCGTGC 242

|||||

Db 88 TTTGGTGVKTTTVNSGWNRCNSVYVWBTAYCDYBHYBRANHVDTTCTNDRGTC 147

|||||

QY 243 CTGGGACGACGCCGCCGCCGCCGCCCTCTTCCGCCAGGTGCTCCTGCTGAAGG 302

|||||

Db 148 NYTASDNGTSATKRVTYDKTSDCGGCKWRKVTYSSBYBRCGVNVMVTTTSMWTDKST 207

|||||

QY 303 AGCTGGTGGCCGAGTGCTGCAGAGGCTGTGCGAGCGCGCGGCGGAGACGCTGCTGCCT 362

|||||

Db 208 KMSMDMSRRSRVHYGRWMBNKRKMSR 235

|||||

QY 363 TCGGCTTCGGGCTGCTGGACGGGGCCCG 390

|||||

RESULT 9

LOCUS AA754458 247 bp mRNA EST 20-JAN-1998

DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa

CDNA clone 97SN1784, mRNA sequence.

ACCESSION AA754458

NID 92801164

KEYWORDS EST.

SOURCE rice.

ORGANISM Oryza sativa

Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;

Polales; Poaceae; Oryza.

1 (bases 1 to 247)

Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,

Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,

Lee,M.C. and Eun,M.Y.

Large-scale Sequencing Analysis of ESTs from Rice Immature Seed

Unpublished (1998)

Contact: Eun M.Y.

Department of Cytogenetics

National Inst. of Agri. Sci. and Tech, RDA

Suwon, Kyunggido, Korea

Tel: 82 331 290 0301

Fax: 82 331 290 0307

Email: myeun@sun20.asti.re.kr

Submitted by Baek Hie Nam, Dept of Biological Science, Myongji

University, Yongin, Korea. 449-728 bhnam@bioserver.myongji.ac.kr

Seq primer: M13 Reverse Primer.

Location/Qualifiers

1..247

/organism="Oryza sativa"

/cultivar="Milyang23"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:

XhoI; Directional cDNA library inserted into lambda ZAPII

vector at 5' end with EcoRI and 3' end with Xho I site."

/db\_xref="taxon:4530"

/clone="97SN1784"

/clone\_lib="Rice Immature Seed Lambda ZAPII cDNA Library"

/tissue\_type="Immature Seed"

/dev\_stage="5 days after pollination"

/lab\_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others

ORIGIN

Query Match 1.2%; Score 47; DB 12; Length 247;

Best Local Similarity 14.8%; Pred. No. 2.25e-32;

Matches 27; Conservative 93; Mismatches 88; Indels 0; Gaps 0;

Db 28 TBWCCVRRYGTITNRKNGKRRTTTNDSCDNDNAHCRITYVWYARSKYGYGTBYISNVND 87

|||||

QY 183 TGCAGCGGGGACCCGGCGCTTTCGCGCGCTGTGGCCAGTGCCTGGTGGCGTGC 242

|||||

Db 88 TTTGGTGVKTTTVNSGWNRCNSVYVWBTAYCDYBHYBRANHVDTTCTNDRGTC 147

|||||

QY 243 CTGGGACGACGCCGCCGCCGCCGCCCTCTTCCGCCAGGTGCTCCTGCTGAAGG 302

|||||

Db 148 NYTASDNGTSATKRVTYDKTSDCGGCKWRKVTYSSBYBRCGVNVMVTTTSMWTDKST 207

|||||

QY 303 AGCTGGTGGCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGGAGACGCTGCTGCCT 362

|||||



```

centromeric to human MHC across the 6p21.2-6p21.3 chromosomal
boundary
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 660)
AUTHORS Tripodis,N. and Ragoussis,J.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Division of Medical and Molecular Genetics,
Guy's Hospital, 7th floor, Guy's Tower, London SE1 9RT, UK
FEATURES             Location/Qualifiers
     source           1..660
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
                     /clone="ntcon6 contig"
                     /tissue_type="fetal brain; fetal liver; adult muscle"
                     /note="similar to CutA"
BASE COUNT   162 a    161 c    173 g    109 t    55 others
ORIGIN
Query Match      0.8%; Score 31; DB 11; Length 660;
Best Local Similarity 15.6%; Pred. No. 1.86e-10;
Matches       7; Conservative 31; Mismatches 7; Indels 0; Gaps 0;

Db 384 WSSRCYRSYGMYHYCYKDKMSCTTSRWRKYKSRSWDCD 428
      :| ::|| ::||:: ||::||:: ||::||:: ||::||:: ||::||:
QY 348 AGAACGTGCTGGCCTTCGGCTTCGGCTCGCTGGACGGGCCGCG 392


RESULT 13
LOCUS A1171394          143 bp      mRNA            EST        07-OCT-1998
DEFINITION EST17350 Normalized rat muscle, Bento Soares Rattus sp. CDNA clone
RMUBK19 3' end, mRNA sequence.
ACCESSION A1171394
NID       G3711434
KEYWORDS  EST.
SOURCE    Rattus sp.
ORGANISM  Rattus sp.
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 143)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
          Kerlavage,A.R. and Adams,M.D.
          Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat
          Gene Index
JOURNAL Unpublished (1998)
COMMENT Other_ESTS: TC52270
Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel.: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
          Location/Qualifiers
     source           1..143
                     /organism="Rattus sp."
                     /note="Organ: muscle; Vector: pT73Pac; Site_1: EcoRI;
                     Site_2: NotI"
                     /db_xref="taxon:10118"
                     /clone="RMUBK19"
                     /clone_lib="Normalized rat muscle, Bento Soares"
BASE COUNT   14 a    49 c    58 g    22 t
ORIGIN
Query Match      0.7%; Score 25; DB 18; Length 143;
Best Local Similarity 76.6%; Pred. No. 1.41e-03;
Matches      36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 4 GGGAGCGCATGGCGCGTGCGGGGCTCGCGCGCGTCCCGCTGGTG 50
|||||
Cc 104 GGGAGCGCATGGCTCGGCAGCGGGAGCGCGGCATCGCGGGGTG 58
|||||

```

LOCUS	AA855630	187 bp	mrna	EST	06-MAR-1998				
DEFINITION	X660115.1 r1 Stratiogene mouse heart (#937316) Mus musculus cDNA clone sequence.								
ACCESSION	J260165.5, similar to TR:Q99960 Q99960 PLAKOPHILIN 2A. [1] ;, MRNA								
NID	AA855630								
KEYWORDS	G2943168								
SOURCE	EST.								
ORGANISM	house mouse.								
	Mus musculus								
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;								
	Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
REFERENCE	1 (bases 1 to 187)								
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.								
TITLE	The WashU-HMI Mouse EST Project								
JOURNAL	Unpublished (1996)								
COMMENT									

Contact: Maria M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

MGI:662717  
Seq primer: -28m13 rev1 ET from Amersham  
High quality sequence stop: 178.

**FEATURES**  
**source**

```

1. 10/
/organism="Mus musculus"
/strain="NIH/Swiss"
/note="Organ: heart; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dr. 93 pooled NIH/Swiss 13 day embryos hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
/db_xref="taxon:10090"
/clone="1260165"
/clone_lib="Stratagene mouse heart (#37316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
29 a 65 c 70 q 23 t

```

BASE COUNT  
ORIGIN

Query Match	0.7%	Score 25;	DB 13;	Length 187;
Best Local Similarity	85.7%;	Pred. No. 1.41e-03;		
Matches	30;	Conservative 0;	Mismatches 5;	Indels 0;
				Caps 0;
Db	105	GGCTGGCGGCAGCAGCGGCGCGGCACCGCGGCGC	139	
Qy	170	GGCTGGCGGCTGGTCAGCGCGGGGACCGCGGCGC	204	

Search completed: Fri Dec 25 05:48:10 1998  
Job time : 13715 secs.

\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm with gap extension penalties as well as gap opening penalties

```
Run on:      Fri Dec 25 05:48:30 1998; MasPar time 1203.77 Seconds
           429.257 Million cell updates/sec

Tabular output not generated.
```

Title: >US-08-951-733-19

Description: (1-3798) from US08951733, seq

perfect score: 3798

N.A. Sequence: 1 CCACGCGTCCGGCGACGCT.....GGAATAGTCATCCCTGAT 3798  
Comp: GGTGCCAGGCCCGTCGCGA.....CCTTATCAGGTAGGGGACTA

Scoring table: TABLE default

Gap open 30: Gap extend 1

Nmatch  
STD : Dbase 0: Query 0

Searched: 188442 seqs. 68026449 bases x 2

Post-processing: Minimum Match 0%

Processing: Minimum match of  
Listing first 45 summaries

Database: n-genes32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 9.981: Variance 6.350: scale 1.572

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			ID	Description	Pred. No.
	Score	Match	Length			
1	45	1.2	91	Q51746	Oligonucleotide probe	2.38e-10
c	45	1.2	91	Q51746	Oligonucleotide probe	2.38e-10
3	40	1.1	114	Q70466	Generic DNA sequence	1.13e-07
c	40	1.1	172	T76363	Human interleukin 8	1.13e-07
5	43	1.1	204	N81154	Base substituted E.co	2.87e-09
c	41	1.1	204	N81154	Base substituted E.co	3.34e-08
7	39	1.0	114	Q70469	Generic DNA sequence	3.78e-07
8	38	1.0	114	Q70468	Generic DNA sequence	1.25e-06
9	38	1.0	114	Q70465	Generic DNA sequence	1.25e-06
c	38	1.0	114	Q70465	Generic DNA sequence	1.25e-06
c	38	1.0	114	Q70468	Generic DNA sequence	1.25e-06
c	38	1.0	114	Q70468	Generic DNA sequence	1.25e-06

C	13	37	1.0	114	12	Q70459	Generic DNA sequence	4.11e-06
	14	39	1.0	178	32	T76405	Human endothelin-1 an	3.78e-07
C	15	38	1.0	178	32	T76405	Human endothelin-1 an	1.25e-06
	16	36	0.9	114	12	Q70467	Generic DNA sequence	1.33e-05
	17	36	0.9	114	12	Q70470	Generic DNA sequence	1.33e-05
	18	34	0.9	114	12	Q70472	Generic DNA sequence	1.36e-04
C	19	36	0.9	114	12	Q70466	Generic DNA sequence	1.33e-05
C	20	36	0.9	114	12	Q70470	Generic DNA sequence	1.33e-05
	21	35	0.9	162	32	T76307	Human RANTES antisense	4.29e-05
	22	34	0.9	168	32	T76270	Human MDNCF antisense	1.36e-04
	23	35	0.9	172	32	T76363	Human interleukin 8 a	4.29e-05
	24	34	0.9	172	32	T76376	Human GM-CSF antisense	1.36e-04
	25	35	0.9	190	32	T76452	Chymase antisense oli	4.29e-05
C	26	35	0.9	190	32	T76452	Chymase antisense oli	4.29e-05
	27	31	0.8	88	32	T76170	Human IL3 receptor an	4.06e-03
C	28	32	0.8	88	32	T76170	Human IL3 receptor an	1.33e-03
	29	30	0.8	100	32	T76186	Human IL4 receptor an	1.22e-02
C	30	30	0.8	100	32	T76186	Human IL4 receptor an	1.22e-02
	31	32	0.8	114	12	Q70471	Generic DNA sequence	1.33e-03
	32	31	0.8	114	12	Q70473	Generic DNA sequence	4.06e-03
C	33	32	0.8	114	12	Q70472	Generic DNA sequence	1.33e-03
C	34	31	0.8	114	12	Q70473	Generic DNA sequence	4.06e-03
C	35	30	0.8	114	12	Q70471	Generic DNA sequence	1.22e-02
	36	32	0.8	128	32	T76233	Human IL6 antisense o	1.33e-03
C	37	31	0.8	128	32	T76233	Human IL6 antisense o	1.33e-03
C	38	30	0.8	130	32	T76152	Human vascular cell a	1.22e-02
	39	29	0.8	150	32	T76238	Human IL6 receptor an	3.63e-02
C	40	29	0.8	168	32	T76270	Human MDNCF antisense	3.63e-02
	41	30	0.8	201	2	N70194	Signal portion of gen	1.22e-02
	42	30	0.8	201	2	N70195	Streptomyces protease	1.22e-02
	43	29	0.8	250	32	T76438	Substance P antisense	3.63e-02
	44	29	0.8	317	32	T76274	Human neutrophil elas	3.63e-02
	45	30	0.8	565	6	Q35072	HCV envelope region p	1.22e-02

## ALIGNMENTS

RESULT	1	
ID	Q51746 standard; cDNA; 91 BP.	
AC	Q51746;	
DT	31-MAY-1994 (first entry)	
DE	Oligonucleotide probe MK14-A	
KW	Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;	
KW	ss.	
OS	Synthetic.	
PN	EP-571911-A.	
PD	01-DEC-1993.	
PF	24-MAY-1993; 108325.	
PR	26-MAY-1992; US-889651.	
PP	(BECT ) BECTON DICKINSON CO.	
PA	Shank DD, Spears PA;	
PI	WPI: 93-378844/48.	
DR	New oligo-nucleotide probes specific for Mycobacteria - used for	
PPT	detection and amplification of Mycobacteria nucleic acid in	
PPT	samples	
PPT	Claim 3; Page 14; 23pp; English.	
PC	Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14	
CC	(Q51735). It hybridized to all spp. of mycobacteria tested, but	
CC	cross reacted to a few non-mycobacterial spp. The probe may	
CC	be useful as an initial screen for mycobacterial infection.	
CC	see also Q51735-45 and Q51747-59.	
CC	sequence 91 BP: 5 A; 17 C; 15 G; 4 T;	
CC		

Query Match	1.2%	Score 45;	DB 9;	Length 91;
Best Local Similarity	7.3%;	Pred. NO. 2.38e-10;		
Matches	4;	Conservative 46;	Mismatches 5;	Indels 0;
Gaps	0;			
Db	6	cgcgsgshvsvvvvshhhshvhhvhhvsvvvvhhvhhvhhvhhvsvv	60	
Qy	391	ggggggccccggagccgttcacacacagcgtgcccactacctgcccaacag	445	

## RESULT





Cp 3427 GTTGGCTGGCGGCTCCAGGCGCATCAGCTGCTGCCCG 3390

RESULT 5

ID	N81164	standard; DNA; 204 BP.
AC	N81164;	
AD	08-NOV-1990	(first entry)
DE	Base substituted E.coli beta-galactosidase alpha-fragment.	
DW	E.coli beta galactosidase alpha-fragment; base substitutions; ss.	
KW	Escherichia coli.	
OS	Escherichia coli.	
FH	Key	Location/Qualifiers
FT	misc_feature	19..69
FT		/tag= a
FT		/function-multiple cloning site
FT	primer_bind	187..204
FT		/*tag= b
FN	EP-285123-A.	
PD	05-MAY-1988.	
PP	30-MAR-1988;	105163.
PP	03-APR-1987;	US-034819.
PA	(SUVO) SUOMEN SOKERI OY.	
PI	Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;	
PI	WPI: 88-279927/40.	
DR	Introducing random point mutations into nucleic acids -	
PT	by prepn of single stranded template, annealing a primer, elongation,	
PT	misincorporation, completion of molecules and screening.	
PS	Disclosure; p; English.	
CC	Random point mutations were introduced into the alpha fragment of	
CC	E.coli beta-galactosidase. The wild type sequence was obtained as a	
CC	single stranded template and an oligonucleotide was hybridised to	
CC	it to generate a popn of DNA molecules which terminate at all	
CC	possible nucleotide positions within a specified region. The	
CC	variable 3' ends generated in this way are used as primers for	
CC	reverse transcriptase. Nucleotides are misincorporated by the	
CC	transcriptase and the molecules are completed to forms that can be	
CC	amplified and then expressed in a suitable host-vector system.	
CC	The sequence covers all 176 diff base substitutions, most of which	
CC	occurred spontaneously in any given mutant.	
CC	See also P80575.	
SQ	Sequence	204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Watch 1.18; Score 43; DB 1; Length 204;

Best Local Similarity 20.9%; Pred. No. 2.87e-09;

Matches 31; Conservative 57; Mismatches 60; Indels 0; Gaps

Db 45 bcyrraggnycgccgaggywccgagcycaaycdchvcgcmrtthhyrrmbnvrd 104

QY 830 TCTGTGGCCACCCGGGACGAGCGGTGGACCGACCTCTTTGGAGGTCGCTCTCTGCGACGCGCAC 899

Db 105 ynrsdaawycyrrsvkycynachdhdyvbbvynvnhnnccebnhchv 164

QY 890 CCTGCCACGACCCCGGAAGAAGCACCTCTTTGGAGGTCGCTCTCTGCGACGCGCAC 919

Db 165 hbnhnrtwayrhdarrddvhcvcchcc 192

QY 950 TCCACCCCATCCGTGGCGCCGACGACC 977

RESULT 6

ID	N81164	standard; DNA; 204 BP.
AC	N81164;	
AD	08-NOV-1990	(first entry)
DE	Base substituted E.coli beta-galactosidase alpha-fragment.	
DW	E.coli beta galactosidase alpha-fragment; base substitutions; ss.	
KW	Escherichia coli.	
OS	Escherichia coli.	
FH	Key	Location/Qualifiers
FT	misc_feature	19..69
FT		/tag= a
FT		/function-multiple cloning site
FT	primer_bind	187..204
FT		/*tag= b
FN	EP-285123-A.	
PD	05-MAY-1988.	









**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*

WORLD

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nna n.a. - n.a. database search, using Smith-Waterman algorithm  
with gap extension penalties as well as gap opening penalties

Run on: Fri Dec 25 06:26:25 1998; MasPar time 6934.12 Seconds  
Tabular output not generated. 586.919 Million cell updates/sec

Title: >US-08-951-733-19  
Description: (1-3798) from US08951733.seq  
Perfect Score: 3798  
N.A. Sequence: 1 CCACGCGTCCGGCAGCGCT.....GGAATAGTCATCCCTGAT 3798  
Comp: GGTGCGCAGGCGCGTCGCGA.....CCTATCAGGTAGGGGACTA

Scoring table: TABLE default  
Gap open 30; Gap extend 1

Nmatch STD : Dbase 0; Query 0

Searched: 1665728 seqs, 535777161 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-pending  
1:P9 2:U6000 3:U6001 4:U6002 5:U6003 6:U6004 7:U6005  
8:U6006 9:U6007 10:U6008 11:U7 12:U80 13:U81 14:U82  
15:U83 16:U84 17:U84B 18:U85 19:U86 20:U87 21:U88  
22:U89 23:U90A 24:U90B 25:U91 26:NEWP 27:NEWU6 28:NEWU8  
29:NEWU9

Statistics: Mean 11.275; Variance 4.495; scale 2.508

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3798	100.0	3798	22	US-08-951-Sequence 19, Applicati	0.00e+00
2	3787	99.7	7029	22	US-08-911-Sequence 1, Applicati	0.00e+00
3	3784	99.6	4015	24	US-09-052-Sequence 1, Applicati	0.00e+00
4	3784	99.6	4015	22	US-08-912-Sequence 1, Applicati	0.00e+00
5	3784	99.6	4015	21	US-08-854-Sequence 224, Applicat	0.00e+00
6	3781	99.6	4023	23	US-09-026-Sequence 35, Applicati	0.00e+00
7	3468	91.3	4029	21	US-08-854-Sequence 173, Applicat	0.00e+00
8	3468	91.3	4029	21	US-08-851-Sequence 173, Applicat	0.00e+00
9	3391	89.3	3855	22	US-08-912-Sequence 4, Applicati	0.00e+00
10	3391	89.3	3855	22	US-08-911-Sequence 18, Applicati	0.00e+00
11	2848	75.0	2848	22	US-08-951-Sequence 13, Applicati	0.00e+00
12	1711	45.1	2176	22	US-08-912-Sequence 3, Applicati	0.00e+00
13	1575	41.5	3346	23	US-09-026-Sequence 36, Applicati	0.00e+00
14	1520	40.0	2171	21	US-08-851-Sequence 100, Applicat	0.00e+00
15	1520	40.0	2171	21	US-08-854-Sequence 100, Applicat	0.00e+00

16	1520	40.0	2171	21	US-08-846-Sequence 100, Applicat	0.00e+00
17	1497	39.4	4200	22	US-08-912-Sequence 6, Applicati	0.00e+00
18	1079	28.4	3496	23	US-09-042-Sequence 1, Applicati	0.00e+00
19	949	25.0	949	22	US-08-951-Sequence 18, Applicati	0.00e+00
20	535	25.0	535	22	US-08-911-Sequence 52, Applicati	0.00e+00
21	460	12.1	550	22	US-08-911-Sequence 54, Applicati	0.00e+00
22	385	10.1	389	21	US-08-846-Sequence 62, Applicati	0.00e+00
23	385	10.1	389	21	US-08-844-Sequence 62, Applicati	0.00e+00
24	385	10.1	389	21	US-08-854-Sequence 62, Applicati	0.00e+00
25	385	10.1	389	22	US-08-912-Sequence 8, Applicati	0.00e+00
26	385	10.1	389	22	US-08-911-Sequence 17, Applicati	0.00e+00
27	385	10.1	389	21	US-08-851-Sequence 62, Applicati	0.00e+00
28	182	4.8	182	22	US-08-912-Sequence 9, Applicati	2.49e-148
29	120	3.2	2651	23	US-09-042-Sequence 7, Applicati	1.04e-85
30	97	2.6	240	22	US-08-912-Sequence 5, Applicati	3.64e-63
31	83	2.2	249	5	US-60-035-Sequence 1054, Applicati	9.44e-50
32	78	2.1	578	18	US-08-569-Sequence 14, Applicati	4.90e-45
33	78	2.1	578	19	US-08-624-Sequence 14, Applicati	4.90e-45
34	69	1.8	477	11	US-07-904-Sequence 10, Applicati	1.12e-36
35	69	1.8	477	11	US-07-904-Sequence 10, Applicati	1.12e-36
36	68	1.8	578	19	US-08-624-Sequence 14, Applicati	9.30e-36
37	68	1.8	578	18	US-08-569-Sequence 14, Applicati	9.30e-36
38	69	1.8	1808	23	US-09-042-Sequence 4, Applicati	1.12e-36
39	67	1.8	7218	17	US-08-466-Sequence 14, Applicati	7.65e-35
40	66	1.7	383	16	US-08-446-Sequence 8, Applicati	6.25e-34
41	66	1.7	383	15	US-08-311-Sequence 8, Applicati	6.25e-34
42	66	1.7	383	16	US-08-446-Sequence 8, Applicati	6.25e-34
43	66	1.7	383	16	US-08-451-Sequence 8, Applicati	6.25e-34
44	66	1.7	383	16	US-08-446-Sequence 8, Applicati	6.25e-34
45	66	1.7	383	16	US-08-446-Sequence 8, Applicati	6.25e-34

ALIGNMENTS

RESULT 1  
ID US-08-951-733-19 STANDARD; DNA; UNC; 3798 BP.  
AC xxxxxx  
DE Sequence 19, Application US/08951733  
Sequence 19, Application US/08951733  
CC GENERAL INFORMATION:  
CC APPLICANT: Harrington, Lea A.  
CC APPLICANT: Robinson, Murray O.  
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
CC NUMBER OF SEQUENCES: 44  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Amgen Inc.  
CC STREET: One Amgen Center Drive  
CC CITY: Thousand Oaks  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 91320-1789  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/951,733  
CC FILING DATE: 16-OCT-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/873,039  
CC FILING DATE: 11-JUN-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/751,189  
CC FILING DATE: 15-NOV-1996  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Oleski, Nancy A.  
CC REGISTRATION NUMBER: 34,688  
CC REFERENCE/DOCKET NUMBER: A-433B  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (805) 447-6504

CC TELEFAX: (805) 499-8011

CC INFORMATION FOR SEQ ID NO: 19:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 3798 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cDNA

SQ SEQUENCE 3798 BP: 613 A; 1310 C; 1213 G; 662 T; 0 OTHER.

Query Match 100.0%; Score 3798; DB 22; Length 3798;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 3798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CCAGCGTCGCGGAGCGCTGCTGCTGCTGCGACAGTGGGAAGCCCTGCGCCGCGCAC 60

QY 1 CCAGCGTCGCGGAGCGCTGCTGCTGCTGCGACAGTGGGAAGCCCTGCGCCGCGCAC 60

Db 61 CCCCGCATGCGCGCGCTCCCGCTCCCGAGCCGCTGCTGCTGCGACCCACTA 120

QY 61 CCCCGCATGCGCGCGCTCCCGCTCCCGAGCCGCTGCTGCTGCGACCCACTA 120

Db 121 CCGCGAGTGTGCGCTGCGCAGCTTCTGCGCGCTGCGCGCCCGAGGCTGGCGGCT 180

QY 121 CCGCGAGTGTGCGCTGCGCAGCTTCTGCGCGCTGCGCGCCCGAGGCTGGCGGCT 180

Db 181 GGTGACGCGGGGACCCGGCGCTTTCGCGCGCTGCTGCGCGCCAGTGCCTGCTGCTG 240

QY 181 GGTGACGCGGGGACCCGGCGCTTTCGCGCGCTGCTGCGCGCCAGTGCCTGCTGCTG 240

Db 241 GCCGTGGAGCAGCG 300

QY 241 GCCGTGGAGCAGCG 300

Db 301 GGAGCTGGTGGCCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGCGAGACGCTGCTGGC 360

QY 301 GGAGCTGGTGGCCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGCGAGACGCTGCTGGC 360

Db 361 CTTGCGCTTGGCGTGTGAGAGGGGCGCGCGGGGCGCGCGGGGCGCGCGCGCGCGCG 420

QY 361 CTTGCGCTTGGCGTGTGAGAGGGGCGCGCGGGGCGCGCGGGGCGCGCGCGCGCGCG 420

Db 421 CGTGGCAGCTACCTGCGCCACACGCTGACCGACGCTGCGGGGAGCGGGGCGTGGG 480

QY 421 CGTGGCAGCTACCTGCGCCACACGCTGACCGACGCTGCGGGGAGCGGGGCGTGGG 480

Db 481 GCTGCTGTGCGCGCTGGCGACGAGCTGTGTTTACCTGCTGCGACGCTGCGCGCT 540

QY 481 GCTGCTGTGCGCGCTGGCGACGAGCTGTGTTTACCTGCTGCGACGCTGCGCGCT 540

Db 541 CTTTGTGCTGCTGCTCCAGCTCGCGCTACCGAGTGTGCGGGCGCGCGCTGTACCACT 600

QY 541 CTTTGTGCTGCTGCTCCAGCTCGCGCTACCGAGTGTGCGGGCGCGCGCTGTACCACT 600

Db 601 CCGGCTGTCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

QY 601 CCGGCTGTCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660

Db 661 ATGCGAAGCGGCTTGAACATAGCGTCAGGAGGCGCGGGTCCCGCTGGCGCTGCCAGC 720

QY 661 ATGCGAAGCGGCTTGAACATAGCGTCAGGAGGCGCGGGTCCCGCTGGCGCTGCCAGC 720

Db 721 CCGGGTGCAGAGGCGCGGGGCGAGTGCAGCGCGAGTGTGCGGTGCGCGAAGAGGCC 780

QY 721 CCGGGTGCAGAGGCGCGGGGCGAGTGCAGCGCGAGTGTGCGGTGCGCGAAGAGGCC 780

Db 781 CAGCGTGGCGCTCCCGCTGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840

QY 781 CAGCGTGGCGCTCCCGCTGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840

Db 841 CCGGGGAGGACGCTGGACGAGTGCAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

QY 841 CCGGGGAGGACGCTGGACGAGTGCAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Db 901 CGCCGAAGAGCCACCTCTTTGGAGGGTGGCTCTCTGGACGCGCCACTCTCCACCACATC 960  
QY 901 CGCCGAAGAGCCACCTCTTTGGAGGGTGGCTCTCTGGACGCGCCACTCTCCACCACATC 960  
Db 961 CGTGGCGCGCCAGCAGCAGCGGCGCCCATCCACATCGGCGCCAGCTGCTCCCTGGGA 1020  
QY 961 CGTGGCGCGCCAGCAGCAGCGGCGCCCATCCACATCGGCGCCAGCTGCTCCCTGGGA 1020  
Db 1021 CAGCGCTTGTCCCGCGGTGTACCGCGAGACCAAGCACTTCTCTACTCTCTCAGCGCACAA 1080  
QY 1021 CAGCGCTTGTCCCGCGGTGTACCGCGAGACCAAGCACTTCTCTACTCTCTCAGCGCACAA 1080  
Db 1081 GGAGCAGCTGCGGCGCTCTCTCTACTCTCTGAGGCGCGAGCTGACTGGCGCTG 1140  
QY 1081 GGAGCAGCTGCGGCGCTCTCTCTACTCTCTGAGGCGCGAGCTGACTGGCGCTG 1140  
Db 1141 GAGGCTGTGGAGACCATCTTCTGGGTTCAGGCGCTTGATGCCAGGACATCCCGCGAG 1200  
QY 1141 GAGGCTGTGGAGACCATCTTCTGGGTTCAGGCGCTTGATGCCAGGACATCCCGCGAG 1200  
Db 1201 GTTGGCGCGCTGCGCGAGCGCTTACTGGCAATGCGGCGCTTCTTCTGGAGCTGCTGG 1260  
QY 1201 GTTGGCGCGCTGCGCGAGCGCTTACTGGCAATGCGGCGCTTCTTCTGGAGCTGCTGG 1260  
Db 1261 GAACACGCGCAGTGTGCGCGCTTACGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTG 1320  
QY 1261 GAACACGCGCAGTGTGCGCGCTTACGGGTGTCTCTCAAGACGCACTGCCCGCTGCGAGCTG 1320  
Db 1321 GGTACCGCGCAGCAGCGGTGTCTGTCGCGGGAGAGCGCGGCTGTGTGGCGCGCC 1380  
QY 1321 GGTACCGCGCAGCAGCGGTGTCTGTCGCGGGAGAGCGCGGCTGTGTGGCGCGCC 1380  
Db 1381 CGAGGAGGAGACACAGACCGCGCTGCTGGTGGAGCTGCTCCGCGCAGCAGCAGCGCC 1440  
QY 1381 CGAGGAGGAGACACAGACCGCGCTGCTGGTGGAGCTGCTCCGCGCAGCAGCAGCGCC 1440  
Db 1441 CTGGCAGGTGTACGGCTTCTGTCGCGGCTGCTGTCGCGGCTGCTGTCGCGGCTGCTG 1500  
QY 1441 CTGGCAGGTGTACGGCTTCTGTCGCGGCTGCTGTCGCGGCTGCTGTCGCGGCTGCTG 1500  
Db 1501 GGCTCCAGGACAGCAGCGCGCTTCTCAGAGACCAAGAAAGTTTCTCTCCCTGGG 1560  
QY 1501 GGCTCCAGGACAGCAGCGCGCTTCTCAGAGACCAAGAAAGTTTCTCTCCCTGGG 1560  
Db 1561 GAAGCATGCCAAGCTCTGCTGCGAGGCTGAGCTGGAAGATGAGCTGCGGAGCTGCGC 1620  
QY 1561 GAAGCATGCCAAGCTCTGCTGCGAGGCTGAGCTGGAAGATGAGCTGCGGAGCTGCGC 1620  
Db 1621 TTGGCTGCGCAGGAGCGCGCTGCTGTTCCGCGCGCAGACCGCTGCTGCTGA 1680  
QY 1621 TTGGCTGCGCAGGAGCGCGCTGCTGTTCCGCGCGCAGACCGCTGCTGCTGA 1680  
Db 1681 GGAGATCTGGCGCAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
QY 1681 GGAGATCTGGCGCAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
Db 1741 GTCTTTCTTTTATGTACGAGACCACTTTTCAAAAGACAGGCTCTTTTCTACCGGAA 1800  
QY 1741 GTCTTTCTTTTATGTACGAGACCACTTTTCAAAAGACAGGCTCTTTTCTACCGGAA 1800  
Db 1801 GAGTGTCTGGAGCAAGTTTGAAGCAATTTGAATTCAGACGACTTGAAGAGGCTGAGCT 1860  
QY 1801 GAGTGTCTGGAGCAAGTTTGAAGCAATTTGAATTCAGACGACTTGAAGAGGCTGAGCT 1860  
Db 1861 GCGGAGCTGTGGAAGCAGAGCTCAGGAGCATCGGAGCGCGCGCGCGCGCGCGCGCG 1920  
QY 1861 GCGGAGCTGTGGAAGCAGAGCTCAGGAGCATCGGAGCGCGCGCGCGCGCGCGCGCG 1920  
Db 1921 GTCCAGACTCCGCTTCTATCCCAAGCTGACGGGCTGACGGGCTGACGGGCTGACGACTA 1980  
QY 1921 GTCCAGACTCCGCTTCTATCCCAAGCTGACGGGCTGACGGGCTGACGGGCTGACGACTA 1980



Db 1981 CGTCGTGGAGCCAGAACGTTCCGACAGAGAAAGAGGCGCGCGTCTCACCTCGAGGCT 2040  
|||||  
QY 1981 CGTCGTGGAGCCAGAACGTTCCGACAGAGAAAGAGGCGCGTCTCACCTCGAGGCT 2040  
|||||  
Db 2041 GAAGGCACTGTTTCAGCGGTGCTCAACTACAGAGCGGCGCGCGCGCTCCCTCGGCGC 2100  
|||||  
QY 2041 GAAGGCACTGTTTCAGCGGTGCTCAACTACAGAGCGGCGCGCGCGCTCCCTCGGCGC 2100  
|||||  
Db 2101 CTCTGCTGGGCGCTGGACGATATCCACAGGCGCTGGCGACCTTCGTGCTGGGTGTGG 2160  
|||||  
QY 2101 CTCTGCTGGGCGCTGGACGATATCCACAGGCGCTGGCGACCTTCGTGCTGGGTGTGG 2160  
|||||  
Db 2161 GGCCCAAGGACCCCGCTGAGCTGTACTTTGTCAAGGTGGATGACGGGCGCTACGA 2220  
|||||  
QY 2161 GGCCCAAGGACCCCGCTGAGCTGTACTTTGTCAAGGTGGATGACGGGCGCTACGA 2220  
|||||  
Db 2221 CACCATCCCCCAGGAGAGGCTCAGGAGGTCAATCGCAGCATCATCAAAACCCAGAAAC 2280  
|||||  
QY 2221 CACCATCCCCCAGGAGAGGCTCAGGAGGTCAATCGCAGCATCATCAAAACCCAGAAAC 2280  
|||||  
Db 2281 GTACTGCGTGGGTGATGCCGTGTGTCAGAGGCGCGCCATGGGACGTCGCGCAAGGC 2340  
|||||  
QY 2281 GTACTGCGTGGGTGATGCCGTGTGTCAGAGGCGCGCCATGGGACGTCGCGCAAGGC 2340  
|||||  
Db 2341 CTTCAAGAGCCAGCTCTCTACCTTGACAGACCTCCAGCGGTATCGGACAGTTCGTGGC 2400  
|||||  
QY 2341 CTTCAAGAGCCAGCTCTCTACCTTGACAGACCTCCAGCGGTATCGGACAGTTCGTGGC 2400  
|||||  
Db 2401 TCACCTGTCAGGAGACACGCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCCTCCCT 2460  
|||||  
QY 2401 TCACCTGTCAGGAGACACGCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCCTCCCT 2460  
|||||  
Db 2461 GAATGAGGCGCAGTGGGCTCTTCGACGCTCTTCCTACGCTTCATGTGCCACACGCGCT 2520  
|||||  
QY 2461 GAATGAGGCGCAGTGGGCTCTTCGACGCTCTTCCTACGCTTCATGTGCCACACGCGCT 2520  
|||||  
Db 2521 GCGCATCAGGGGCAAGTCTCTACGTCAGTCCAGGGGATCCCGAGGCTCCATCTCTC 2580  
|||||  
QY 2521 GCGCATCAGGGGCAAGTCTCTACGTCAGTCCAGGGGATCCCGAGGCTCCATCTCTC 2580  
|||||  
Db 2581 CAGCTGCTCTGACGCTGTGCTACGCGACATGGAGAAACAGCTGTTTCGGGGATTCG 2640  
|||||  
QY 2581 CAGCTGCTCTGACGCTGTGCTACGCGACATGGAGAAACAGCTGTTTCGGGGATTCG 2640  
|||||  
Db 2641 GCGGAGCGGCTCTCTCGGTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2700  
|||||  
QY 2641 GCGGAGCGGCTCTCTCGGTTTGGTGATGATTTCTTTGGTGACACCTCACCTCAC 2700  
|||||  
Db 2701 CCAGCGGAAACCTTCTCAGGACCTGTGTCAGGCTGTCGAGGTGCCCTGAGTGGCTGGT 2760  
|||||  
QY 2701 CCAGCGGAAACCTTCTCAGGACCTGTGTCAGGCTGTCGAGGTGCCCTGAGTGGCTGGT 2760  
|||||  
Db 2761 GAACTTCGGAAGACAGTGTGTAACCTTCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820  
|||||  
QY 2761 GAACTTCGGAAGACAGTGTGTAACCTTCCTGTAGAGAGAGGCGCTGGGTGGCAGGC 2820  
|||||  
Db 2821 TTTTGTTCAGATCCGSCCACCAGGCTTATTCCTGTGGTGGGCTGTGCTGGATACCCG 2880  
|||||  
QY 2821 TTTTGTTCAGATCCGSCCACCAGGCTTATTCCTGTGGTGGGCTGTGCTGGATACCCG 2880  
|||||  
Db 2881 GACCTGGAGGTGACAGGCACTTCTCAGCTATGCCGGAACCTTCATCAGAGCCAGTCT 2940  
|||||  
QY 2881 GACCTGGAGGTGACAGGCACTTCTCAGCTATGCCGGAACCTTCATCAGAGCCAGTCT 2940  
|||||  
Db 2941 CACCTCAACCGGCTTCAAGGCTGGGAGAACATGCTGCGCAAACTCTTTGGGCTTT 3000  
|||||  
QY 2941 CACCTCAACCGGCTTCAAGGCTGGGAGAACATGCTGCGCAAACTCTTTGGGCTTT 3000  
|||||  
Db 3001 GCGGCTGAAGTGCACAGCTGTTTCTGGATTTCAGGTGAACAGCTCCAGAGCGGTG 3060  
|||||  
QY 3001 GCGGCTGAAGTGCACAGCTGTTTCTGGATTTCAGGTGAACAGCTCCAGAGCGGTG 3060  
|||||  
Db 3061 CACCAACATCTACAAGATCCTCTGCTGACAGGCTACAGGTTTCACGCAATGTGTGCTGA 3120

QY 3061 CACCAACATCTACAAGATCCTCTGCTGACGGGTACAGGTTTCACGCAATGTGTGCTGA...3120  
|||||  
Db 3121 GCTCCCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGGTCTCATCTCTGA 3180  
|||||  
QY 3121 GCTCCCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGGTCTCATCTCTGA 3180  
|||||  
Db 3181 CAGGGCTCCCTCTGCTACTCTCATCTCTGAAAGCAAGACGAGGATGTCGTGGGGC 3240  
|||||  
QY 3181 CAGGGCTCCCTCTGCTACTCTCATCTCTGAAAGCCAGAAACGAGGATGTCGTGGGGC 3240  
|||||  
Db 3241 CAAGGCGCGCGCGGCTCTGCCCTCCGAGCGGTGCAAGTGTGTCACCAAGCAT 3300  
|||||  
QY 3241 CAAGGCGCGCGCGGCTCTGCCCTCCGAGCGGTGCAAGTGTGTCACCAAGCAT 3300  
|||||  
Db 3301 CTTGCTCAAGCTGACTTCGACACCGTGTCACTACGTGCCACCTCTGGGTCTACTCAGGAC 3360  
|||||  
QY 3301 CTTGCTCAAGCTGACTTCGACACCGTGTCACTACGTGCCACCTCTGGGTCTACTCAGGAC 3360  
|||||  
Db 3361 AGCCAGAGCGAGTGTGAGTGGAAAGTCCGGGAGCGAGTGTGCTGGAGGCGC 3420  
|||||  
QY 3361 AGCCAGAGCGAGTGTGAGTGGAAAGTCCGGGAGCGAGTGTGCTGGAGGCGC 3420  
|||||  
Db 3421 AGCCAGCGGCGCTGCTCAGACTTCAGACCTTCAAGACCATCTGGACTGATGGCCACCCGCCA 3480  
|||||  
QY 3421 AGCCAGCGGCGCTGCTCAGACTTCAGACCTTCAAGACCATCTGGACTGATGGCCACCCGCCA 3480  
|||||  
Db 3481 CAGCCAGGCGGAGAGCAGACACAGCAGCCCTGTCTACGCGGGCTCTAGCTCCAGGGAG 3540  
|||||  
QY 3481 CAGCCAGGCGGAGAGCAGACACAGCAGCCCTGTCTACGCGGGCTCTAGCTCCAGGGAG 3540  
|||||  
Db 3541 GGAGGGCGGCGCCACACCCAGGCGCGCAGCGTGGAGTCTGAGGCTGAGTGTGTTT 3600  
|||||  
QY 3541 GGAGGGCGGCGCCACACCCAGGCGCGCAGCGTGGAGTCTGAGGCTGAGTGTGTTT 3600  
|||||  
Db 3601 GGCGGAGGCTGCTGATGCTGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGAGTGTG 3660  
|||||  
QY 3601 GGCGGAGGCTGCTGATGCTGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGAGTGTG 3660  
|||||  
Db 3661 CAGCCAGGCTGAGTGTCCAGCACACCTGCTCTACTTCCACAGGCTGGGGCTC 3720  
|||||  
QY 3661 CAGCCAGGCTGAGTGTCCAGCACACCTGCTCTACTTCCACAGGCTGGGGCTC 3720  
|||||  
Db 3721 GGCTCACCCCGGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCCACATAG 3780  
|||||  
QY 3721 GGCTCACCCCGGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCCACATAG 3780  
|||||  
Db 3781 AATAGTCCATCCCTGAT 3798  
|||||  
QY 3781 AATAGTCCATCCCTGAT 3798  
|||||

## RESULT 2

ID US-08-911-312-1 STANDARD; DNA; UNC; 7029 BP.

AC xxxxxx

DT

Sequence 1, Application US/08911312

Sequence 1, Application US/08911312

GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin B.

CC APPLICANT: Andrews, William

CC TITLE OF INVENTION: Telomerase Reverse Transcriptase

CC NUMBER OF SEQUENCES: 170

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Crew LLP

CC STREET: Two Embarcadero Center, Eighth Floor

CC CITY: San Francisco

CC STATE: California





QY 3484 CCAGGCGGAGGAGACACACAGAGCCCTGTACCGCGGGCTCTACCTCCAGGGAGGA 3543  
DB 4258 GGGCGGCGCCACACCCAGGCGCCGACCGCTGGGAGTCTGAGGCTGAGTGTTTGGC 4317  
QY 3544 GGGCGGCGCCACACCCAGGCGCCGACCGCTGGGAGTCTGAGGCTGAGTGTTTGGC 3603  
DB 4318 CGAGGCTGTATGTCGGCTGAAGGCTGAGTGTCCGCTGAGGCTGAGGCTGAGTGTCCAG 4377  
QY 3604 CGAGGCTGTATGTCGGCTGAAGGCTGAGTGTCCGCTGAGGCTGAGGCTGAGTGTCCAG 3663  
DB 4378 CCAAGGCTGAGTGTCCAGACACCTCCGCTTCTCACTTCCACAGGCTGGGCTCGGC 4437  
QY 3664 CCAAGGCTGAGTGTCCAGACACCTCCGCTTCTCACTTCCACAGGCTGGGCTCGGC 3723  
DB 4438 TCCACCCAGGCGGAGCTTTCTCCACAGGAGCCCGGCTTCCACTCCCCACATAGGAAT 4497  
QY 3724 TCCACCCAGGCGGAGCTTTCTCCACAGGAGCCCGGCTTCCACTCCCCACATAGGAAT 3783  
DB 4498 AGTCCATCCCGAT 4512  
QY 3784 AGTCCATCCCGAT 3798

RESULT 3  
ID US-09-052-919-1 STANDARD; DNR; UNC; 4015 BP.  
AC xxxxxx  
DE Sequence 1, Application US/09052919  
DT Sequence 1, Application US/09052919  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin B.  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Antisense Compositions for Detecting and  
CC TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
CC NUMBER OF SEQUENCES: 72  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, Eighth Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94111-3834  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/052,919  
CC FILING DATE: 31-MAR-1998  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/911,312  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/912,951  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/915,503  
CC FILING DATE: 14-AUG-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,549  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/974,584  
CC FILING DATE: 19-NOV-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17618  
CC FILING DATE: 01-OCT-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/US97/17885  
CC FILING DATE: 01-OCT-1997  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Parent, Annette S.  
CC REGISTRATION NUMBER: 42,058  
CC REFERENCE/DOCKET NUMBER: 015389-0036000US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4015 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 56..3454  
CC OTHER INFORMATION: /product= "human telomerase reverse  
CC OTHER INFORMATION: transcriptase (hTRT)"  
CC SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.  
SQ

Query Match 99.6%; Score 3784; DB 24; Length 4015;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 3785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
DB 1 GCAGCGCTGCTGCTGCTGCGACGCTGGGAAGCCCTGGCGCCGCGCCGACCCCGCGATGCC 60  
QY 13 GCAGCGCTGCTGCTGCTGCGACGCTGGGAAGCCCTGGCGCCGCGCCGACCCCGCGATGCC 72  
DB 61 GCGCGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 120  
QY 73 GCGCGCTCCCGCTGCGGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAGGTGCT 132  
DB 121 GCGCGTGGCCACGTTGCTGCGGGGCTGGGGCCCGCAGGCTGGCGCTGGTGCAGGCGGG 180  
QY 133 GCGCGTGGCCACGTTGCTGCGGGGCTGGGGCCCGCAGGCTGGCGCTGGTGCAGGCGGG 192  
DB 181 GGACCCGCGCGCTTTCCGCGCGCTGGTGCCCGAGTGGTGGTGGTGGTGGTGGTGGTGG 240  
QY 193 GGACCCGCGCGCTTTCCGCGCGCTGGTGCCCGAGTGGTGGTGGTGGTGGTGGTGGTGG 252  
DB 241 ACGGCGCGCGCGCGCGCGCGCGCTTCCCTTCCGCGCAGGTGCTTCCCTGAAGGAGGTGGTGGC 300  
QY 253 ACGGCGCGCGCGCGCGCGCGCGCTTCCCTTCCGCGCAGGTGCTTCCCTGAAGGAGGTGGTGGC 312  
DB 301 CCGAGTGTGTCAGAGGCTGTGAGCGCGCGCGCGGAGAACGTCGTGGCTTCGGCTTCGC 360  
QY 313 CCGAGTGTGTCAGAGGCTGTGAGCGCGCGCGCGGAGAACGTCGTGGCTTCGGCTTCGC 372  
DB 361 GCTGCTGGACGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCG 420  
QY 373 GCTGCTGGACGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCG 432



QY 2593 CAGCGTGTCTACGGGACATGGAGAACAAAGCTGTTTCGGGGATTCTGGCGGGACGGGCT 2652  
Db 2641 GCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGGGAAAC 2700  
QY 2653 GCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCACGGGAAAC 2712  
Db 2701 CTTCTCAGGACCTGGTCCGAGGTGCTCCCTGAGTATGGTCCGTGGTGAATTCGGGAA 2760  
QY 2713 CTTCTCAGGACCTGGTCCGAGGTGCTCCCTGAGTATGGTCCGTGGTGAATTCGGGAA 2772  
Db 2761 GACAGTGTGAATTCCTCTGTAGAGAGAGAGCCCTGGGTGACAGCGGCTTTTGTTCAGAT 2820  
QY 2773 GACAGTGTGAATTCCTCTGTAGAGAGAGAGCCCTGGGTGACAGCGGCTTTTGTTCAGAT 2832  
Db 2821 GCCGGCCACAGGCTTATCCCTGTGTGGGCTGTCTGTGATACCCGAGACCTTGGAGGT 2880  
QY 2833 GCCGGCCACAGGCTTATCCCTGTGTGGGCTGTCTGTGATACCCGAGACCTTGGAGGT 2892  
Db 2881 GCAGAGCCACTACTCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCAGCTTCAACCG 2940  
QY 2893 GCAGAGCCACTACTCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCAGCTTCAACCG 2952  
Db 2941 CGGCTTCAAGGCTGGAGGAAACATCGCTGCAAACTCTTTGGGGTCTTGGGCTGAAGTG 3000  
QY 2953 CGGCTTCAAGGCTGGAGGAAACATCGCTGCAAACTCTTTGGGGTCTTGGGCTGAAGTG 3012  
Db 3001 TCACAGCCTGTTTCGAGTGTGAGGTAAGAGCCTCCAGAGCGTGTGCACCAACATCTA 3060  
QY 3013 TCACAGCCTGTTTCGAGTGTGAGGTAAGAGCCTCCAGAGCGTGTGCACCAACATCTA 3072  
Db 3061 CAAGATCTCTCTGCTGCAGGCGTACAGTTTCACGCATGTGTGTCAGCTCCCATTTCA 3120  
QY 3073 CAAGATCTCTCTGCTGCAGGCGTACAGTTTCACGCATGTGTGTCAGCTCCCATTTCA 3132  
Db 3121 TCAGCAAGTGTGGAAGAACCCACATTTTCTGGCGCTCATCTCTGACAGCGCTCCCT 3180  
QY 3133 TCAGCAAGTGTGGAAGAACCCACATTTTCTGGCGCTCATCTCTGACAGCGCTCCCT 3192  
Db 3181 CTGCTACTCCTATCTCTGAAAGCAAGAGCAGGATGTCTGTTGGGGCCAAAGGCGCGC 3240  
QY 3193 CTGCTACTCCTATCTCTGAAAGCAAGAGCAGGATGTCTGTTGGGGCCAAAGGCGCGC 3252  
Db 3241 CGGCGCTCTGCCCTCGAGGCGGTGACGTGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
QY 3253 CGGCGCTCTGCCCTCGAGGCGGTGACGTGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3312  
Db 3301 GACTCGACACCGTGTACCTAGTGGCCACTCTGGGGTCACTCAGGAGAGCCAGAGCSA 3360  
QY 3313 GACTCGACACCGTGTACCTAGTGGCCACTCTGGGGTCACTCAGGAGAGCCAGAGCSA 3372  
Db 3361 GCTAGTCCGAGCTCCCGGGAGACGCTGACTGCCCTGGAGGCGCGAGCCAAACCGGC 3420  
QY 3373 GCTAGTCCGAGCTCCCGGGAGACGCTGACTGCCCTGGAGGCGCGAGCCAAACCGGC 3432  
Db 3421 ACTGCCCTCAGACTCAGACATCTCGACTGTGATGGCCACCGCCACAGCCAGGCGGA 3480  
QY 3433 ACTGCCCTCAGACTCAGACATCTCGACTGTGATGGCCACCGCCACAGCCAGGCGGA 3492  
Db 3481 GAGCAGACACAGCAGCCCTGTACGCGGGCTCTACCTCCAGGAGGAGGGCGGCC 3540  
QY 3493 GAGCAGACACAGCAGCCCTGTACGCGGGCTCTACCTCCAGGAGGAGGGCGGCC 3552  
Db 3541 CACACCCAGGCGCCAGCGCTGGAGTCTGAGGCTGTAGTGTGTCGGCGAGGCGCTG 3600  
QY 3553 CACACCCAGGCGCCAGCGCTGGAGTCTGAGGCTGTAGTGTGTCGGCGAGGCGCTG 3612  
Db 3601 CATGTCCGGCTGAAGGCTGTAGTGTCCGCTGAGGCTGAGGCTGTCCAGCCAGGGCT 3660  
QY 3613 CATGTCCGGCTGAAGGCTGTAGTGTCCGCTGAGGCTGAGGCTGTCCAGCCAGGGCT 3672  
Db 3661 GAGTGTCCAGCACACTCTCGCTTCTACTTCCACAGGCTGGCGCTCGGTCCACCCCA 3720  
QY 3673 GAGTGTCCAGCACACTCTCGCTTCTACTTCCACAGGCTGGCGCTCGGTCCACCCCA 3732

Db 3721 GGGCCAGCTTTTCTCTACCAGGAGCCGGCTTCCACTCCCCACATAGATAGTCCATCC 3780  
QY 3733 GGGCCAGCTTTTCTCTACCAGGAGCCGGCTTCCACTCCCCACATAGATAGTCCATCC 3792  
Db 3781 CCAGAT 3786  
QY 3793 CCTGAT 3798

## RESULT 4

ID US-08-912-951-1 STANDARD; DNA; UNC; 4015 BP.  
AC xxxxxx

DT Sequence 1, Application US/08912951

CC Sequence 1, Application US/08912951

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC TITLE OF INVENTION: THERAPEUTIC METHODS

CC NUMBER OF SEQUENCES: 335

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, 8th Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: United States of America

CC ZIP: 94111

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/912,951

CC FILING DATE: 14-AUG-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/854,050

CC FILING DATE: 09-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/851,843

CC FILING DATE: 06-MAY-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/844,419

CC FILING DATE: 18-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643

CC FILING DATE: 01-OCT-1996

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Apple, Randolph T.

CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002600US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 576-0200

CC TELEFAX: (415) 576-0300

CC INFORMATION FOR SEQ ID NO: 1:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 4015 base pairs



```

CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS 3454
CC LOCATION: 56..3454
CC OTHER INFORMATION: /product= "hprt"
CC OTHER INFORMATION: /note= "human telomerase reverse
CC OTHER INFORMATION: transcriptase (hTERT) catalytic protein
CC OTHER INFORMATION: component"
CC OTHER INFORMATION:
CC SEQUENCE 4015 bp: 663 A; 1363 C; 1375 G; 714 T; 0 OTHER.
SQ

```

|||||  
QY 1933 CTTATCCCAAGCTGACGGGCTCGGCCGATTTGTGAACATGACTACGTCTGTTGGAGC 1992  
Db 1981 CAGAACGTTCCGACAGAAAAGAGGCCGAGCGTCTACCTCGAGGGTGAAGGCACCTGTT 2040  
QY 1993 CAGAACGTTCCGACAGAAAAGAGGCCGAGCGTCTACCTCGAGGGTGAAGGCACCTGTT 2052  
Db 2041 CAGCGTGTCTAACTACGAGGGGCGGGCGGCCCGCCGCTCTCTGGCGCTCTGTGCTGGG 2100  
QY 2053 CAGCGTGTCTAACTACGAGGGGCGGGCGGCCCGCCGCTCTCTGGCGCTCTGTGCTGGG 2112  
Db 2101 CTTGACCATATCCACAGGGCTGCGCACCTCTCGTGTGCTGCTGCGGGGCCAGGACCC 2160  
QY 2113 CTTGACCATATCCACAGGGCTGCGCACCTCTCGTGTGCTGCTGCGGGGCCAGGACCC 2172  
Db 2161 GCGCGCTGAGGTGTATTTGTCAAGGTGGATGTGAGGGCGCGGTACGACACCATCCCCCA 2220  
QY 2173 GCGCGCTGAGGTGTATTTGTCAAGGTGGATGTGAGGGCGCGGTACGACACCATCCCCCA 2232  
Db 2221 GGACAGGCTCAGGAGGTCTACGCGAGCATCATCAAAACCCAGAACACGTACTGGGTGG 2280  
QY 2233 GGACAGGCTCAGGAGGTCTACGCGAGCATCATCAAAACCCAGAACACGTACTGGGTGG 2292  
Db 2281 TCGGTATGCCGTGTCAGAGGCCGCCATGGGCACGTCGCGAAGGCTTCAAGAGCCA 2340  
QY 2293 TCGGTATGCCGTGTCAGAGGCCGCCATGGGCACGTCGCGAAGGCTTCAAGAGCCA 2352  
Db 2341 GGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGGGTCACTTCAGGA 2400  
QY 2353 GGTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGGGTCACTTCAGGA 2412  
Db 2401 GACAGCCCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCCCTGATGAGGCCAG 2460  
QY 2413 GACAGCCCGCTGAGGATGCGGTGCTCATCGAGCAGAGCTCTCCCTGATGAGGCCAG 2472  
Db 2461 CAGTGGCCTCTTCGACGCTTCTCTACGCTTCATGTGTCACACGCGCGCATCAGGG 2520  
QY 2473 CAGTGGCCTCTTCGACGCTTCTCTACGCTTCATGTGTCACACGCGCGCATCAGGG 2532  
Db 2521 CAACTCTACCTGACAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCAGGCTGCTCTG 2580  
QY 2533 CAACTCTACCTGACAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCAGGCTGCTCTG 2592  
Db 2581 CAGCGCTGTCTACGCGACATGGAACAAAGCTGTTTGGGGGATTCGGGGGAGCGGCT 2640  
QY 2593 CAGCGCTGTCTACGCGACATGGAACAAAGCTGTTTGGGGGATTCGGGGGAGCGGCT 2652  
Db 2641 GCTCTCGCTTGTGGATGATTTCTTGTGTGACACTCACCTCACCCAGCGAANAAC 2700  
QY 2653 GCTCTCGCTTGTGGATGATTTCTTGTGTGACACTCACCTCACCCAGCGAANAAC 2712  
Db 2701 CTTCTCAGGACCTTGTCCGAGGTGTCCCTGAGTATGGCTGCGTGTGAACCTTCGGAA 2760  
QY 2713 CTTCTCAGGACCTTGTCCGAGGTGTCCCTGAGTATGGCTGCGTGTGAACCTTCGGAA 2772  
Db 2761 GACAGTGTGAATCTTCCCTGTAGAGACAGAGCCCTGGGTGGCAGCGGCTTTGTTTCAGAT 2820  
QY 2773 GACAGTGTGAATCTTCCCTGTAGAGACAGAGCCCTGGGTGGCAGCGGCTTTGTTTCAGAT 2832  
Db 2821 GCGGCCACAGCCCTATTCCCTGTGTGGGCTGTCTGTGATACCGGACCTGGAGGT 2880  
QY 2833 GCGGCCACAGCCCTATTCCCTGTGTGGGCTGTCTGTGATACCGGACCTGGAGGT 2892  
Db 2881 GCGAGCGACTACTCCAGTATGCCCCGACCTCCATCATAGAGCAGTCTCACTTCAACCG 2940  
QY 2893 GCGAGCGACTACTCCAGTATGCCCCGACCTCCATCATAGAGCAGTCTCACTTCAACCG 2952  
Db 2941 GCGCTTCAAGGCTGGAGAACATCGCTCGAANAACCTTTGGGGTCTTTGGGGCTGAAGTG 3000  
QY 2953 GCGCTTCAAGGCTGGAGAACATCGCTCGAANAACCTTTGGGGTCTTTGGGGCTGAAGTG 3012  
Db 3001 TCACAGCCTGTTTCTGGATTTTCAGTGAACAGCCTCCAGAGGCTGCACCAACATCTA 3060  
|||||

QY 3013 TCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA 3072  
Db 3061 CAAGATCCTCTGCTGAGGCGTACAGGTTTACGCAATGTGTCTGAGCTCCCATTTCA 3120  
QY 3073 CAAGATCCTCTGCTGAGGCGTACAGGTTTACGCAATGTGTCTGAGCTCCCATTTCA 3132  
Db 3121 TCACAAAGTTTGAAGAACCCACATTTTCTGCGGGTCTATCTCTACAGCGGCTCCCT 3180  
QY 3133 TCACAAAGTTTGAAGAACCCACATTTTCTGCGGGTCTATCTCTGACAGCGCTCCCT 3192  
Db 3181 CTGCTACTCATCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCAAAGGCGCGCG 3240  
QY 3193 CTGCTACTCATCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCAAAGGCGCGCG 3252  
Db 3241 CGGCCCTCTGCCCTCCAGAGCCGTCAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT 3300  
QY 3253 CGGCCCTCTGCCCTCCAGAGCCGTCAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT 3312  
Db 3301 GACTCGACACGCTGTCACCTACGTGCCACTCTCTGGGGTCTACTCAGGACAGCCAGAGCGCA 3360  
QY 3313 GACTCGACACGCTGTCACCTACGTGCCACTCTCTGGGGTCTACTCAGGACAGCCAGAGCGCA 3372  
Db 3361 GCTGAGTCGGAAGCTCCCGGGGACGCGTGACTGCCCTGGAGCCGCGACCAACCCGCG 3420  
QY 3373 GCTGAGTCGGAAGCTCCCGGGGACGCGTGACTGCCCTGGAGCCGCGACCAACCCGCG 3432  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGTCACCCGCGCCAGCCAGGCGCGA 3480  
QY 3433 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGTCACCCGCGCCAGCCAGGCGCGA 3492  
Db 3481 GAGCAGACACAGAGCCCTGTACGCGGGGCTCTAGTCTCCAGGAGGAGGCGCGCGCG 3540  
QY 3493 GAGCAGACACAGAGCCCTGTACGCGGGGCTCTAGTCTCCAGGAGGAGGCGCGCGCG 3552  
Db 3541 CACACCCAGGCGCCGCGCTGGAGTCTGAGGCTCAGTGAAGTGTGTCGCGAGGCGCTG 3600  
QY 3553 CACACCCAGGCGCCGCGCTGGAGTCTGAGGCTCAGTGAAGTGTGTCGCGAGGCGCTG 3612  
Db 3601 CATCTCGGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3660  
QY 3613 CATCTCGGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3672  
Db 3661 GAGTGTCCAGCAGACCTGCGGCTTCTCACTTCCCGCAGGCTGGGCTGGGCTCCACCCCA 3720  
QY 3673 GAGTGTCCAGCAGACCTGCGGCTTCTCACTTCCCGCAGGCTGGGCTGGGCTCCACCCCA 3732  
Db 3721 GGGCCAGCTTTTCTCACCAGGAGCCGCTTCCACTCCCATAGTAATAGTCCATCC 3780  
QY 3733 GGGCCAGCTTTTCTCACCAGGAGCCGCTTCCACTCCCATAGTAATAGTCCATCC 3792  
Db 3781 CCAGAT 3786  
QY 3793 CCTGAT 3798  
|||||

## RESULT 5

ID US-08-854-050-224 STANDARD; DNA; UNC; 4015 BP.  
AC xxxxxx

Sequence 224, Application US/08854050

Sequence 224, Application US/08854050

## GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC TITLE OF INVENTION: Novel Telomerase

CC NUMBER OF SEQUENCES: 225

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Crew LLP



STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536

OTHER INFORMATION: transcriptase (hTERT) catalytic protein  
OTHER INFORMATION: component"  
SEQUENCE 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 OTHER.

```
Query Match      99.6%; Score 3784; DB 21; Length 4015;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 3785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Db	1	GCAGCGCTGGCTCTCTGTCGCGACGTGGGAAGCCCTGCGCCCCCGCCACCCCGCGATGCC	60
Qy	13	GCAGCGTGGGTCTCTGTCGCGACGTGGGAAGCCCTGCGCCCCCGCCACCCCGCGATGCC	72
Db	61	GGCGGCTCCCGCTGCGGAGCGCTGGCGTCCCTGCTGCGGAGCCACTACCGCGAGTGCT	120
Qy	73	GGCGGCTCCCGCTGCGGAGCGCTGGCGTCCCTGCTGCGGAGCCACTACCGCGAGTGCT	132
Db	121	GCGCGTGGCCACGTTCTGTGCGGGCGCTTGGGGCCCCAGGCGTGCGGCTGCTGACGCGCG	180
Qy	133	GCGCGTGGCCACGTTCTGTGCGGGCGCTTGGGGCCCCAGGCGTGCGGCTGCTGACGCGCG	192
Db	181	GGACCCGGCGGCTTTTCGCGCGCTGTGTGGCCCAAGTGCCTGTGGTGTGCTTGGGACGC	240
Qy	193	GGACCCGGCGGCTTTTCGCGCGCTGTGTGGCCCAAGTGCCTGTGGTGTGCTTGGGACGC	252

|||||  
QY 1333 AGCGGTGCTGTGCCGAGAGAGCCCAAGGCTCTGTGGCGCCCCCAGGAGGAGGA 1392  
Db 1381 CACAGACCCCGTGCCTGTGTGAGCTGCTCCGCCACACAGAGCCCGCTGGCAGGTGTA 1440  
QY 1393 CACAGACCCCGTGCCTGTGTGAGCTGCTCCGCCACACAGAGCCCGCTGGCAGGTGTA 1452  
Db 1441 CGGTTGCTGGGGCTGCTGCTGCGCGGCTGTGTGCCCGCCAGGCGCTCTGGGGCTCCAGGCA 1500  
QY 1453 CGGTTGCTGGGGCTGCTGCTGCGCGGCTGTGTGCCCGCCAGGCGCTCTGGGGCTCCAGGCA 1512  
Db 1501 CAAGAACCCCGTTCCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAA 1560  
QY 1513 CAAGAACCCCGTTCCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAA 1572  
Db 1561 GCTCTGCTCAGAGAGTACGTGGAGATGAGCGTGGGGACTCGCGCTTGGCTGGCGAG 1620  
QY 1573 GCTCTGCTCAGAGAGTACGTGGAGATGAGCGTGGGGACTCGCGCTTGGCTGGCGAG 1632  
Db 1621 GAGCCAGGGTGGCTGTGTTCCGGCCGAGAGACCGCTCTCGGTGAGAGATCTCTGGC 1680  
QY 1633 GAGCCAGGGTGGCTGTGTTCCGGCCGAGAGACCGCTCTCGGTGAGAGATCTCTGGC 1692  
Db 1681 CAAGTTCTGCTGCTGCTGATGAGTGTACGTGCTGCGAGCTGCTCAGGCTCTTTCTTTTA 1740  
QY 1693 CAAGTTCTGCTGCTGCTGATGAGTGTACGTGCTGCGAGCTGCTCAGGCTCTTTCTTTTA 1752  
Db 1741 TGTACGAGAGACCACTTTTCAAGAACAGCGCTCTTTTCTACCGGAAGAGTCTCTGGAG 1800  
QY 1753 TGTACGAGAGACCACTTTTCAAGAACAGCGCTCTTTTCTACCGGAAGAGTCTCTGGAG 1812  
Db 1801 CAAGTTCAAGAGATTTGGAATCAGACAGCACTTTGAGAGGGTGCAGCTGGCGAGCTGTC 1860  
QY 1813 CAAGTTCAAGAGATTTGGAATCAGACAGCACTTTGAGAGGGTGCAGCTGGCGAGCTGTC 1872  
Db 1861 GGAAGCAGAGTTCAGGAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTCCG 1920  
QY 1873 GGAAGCAGAGTTCAGGAGCATCGGGAAGCCAGGCGCCGCTGCTGACGTCAGACTCCG 1932  
Db 1921 CTTATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACCTCGTGGGAGC 1980  
QY 1933 CTTATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACCTCGTGGGAGC 1992  
Db 1981 CAGAACCTTCGAGAGAAAGAGGCGGAGCGTCTCACCTCAGAGGTGAAGGACTGTT 2040  
QY 1993 CAGAACCTTCGAGAGAAAGAGGCGGAGCGTCTCACCTCAGAGGTGAAGGACTGTT 2052  
Db 2041 CAGGCTGCTCAACTACGAGCGGCGCGGCGCCCGCTCCTGGCGGCTCTGTGTGG 2100  
QY 2053 CAGGCTGCTCAACTACGAGCGGCGCGGCGCCCGCTCCTGGCGGCTCTGTGTGG 2112  
Db 2101 CTTGAGAGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGCTGCGGCCACGAGACCC 2160  
QY 2113 CTTGAGAGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGCTGCGGCCACGAGACCC 2172  
Db 2161 GCGCGCTGAGCTGACTTTGTCAGGTGGATGTGAGGGGCGGCTAGACACCATCCGCCCA 2220  
QY 2173 GCGCGCTGAGCTGACTTTGTCAGGTGGATGTGAGGGGCGGCTAGACACCATCCGCCCA 2232  
Db 2221 GGACAGGCTCACGAGGTATCCGACATCATCAACCCCAAGACAGCTACTGCTGG 2280  
QY 2233 GGACAGGCTCACGAGGTATCCGACATCATCAACCCCAAGACAGCTACTGCTGG 2292  
Db 2281 TCGGTATCGCGTGGTCCAGAAAGCCCGCCATGGCGACGCTCCGCAAGGCTTTCAAGAGCCA 2340  
QY 2293 TCGGTATCGCGTGGTCCAGAAAGCCCGCCATGGCGACGCTCCGCAAGGCTTTCAAGAGCCA 2352  
Db 2341 CGTCTTACCTTGACAGACCTCCAGCGGTACATCGGACAGTTCTGGGCTCACCTGCAGGA 2400  
QY 2353 CGTCTTACCTTGACAGACCTCCAGCGGTACATCGGACAGTTCTGGGCTCACCTGCAGGA 2412  
Db 2401 GATCAGCCGCTGAGGATGCCGTGTCATCGAGCAGAGCTCTCCCTGATCAGGCCAG 2460  
|||||

QY 2413 GACCAGCCCGCTGAGGGATGCCGTGCTATCGAGCAGAGACTCCTCCCTGAATGAGGCCAG 2472  
Db 2461 CAGTGGCCCTTCGAGCGTCTTCTACGCTTTCATGTGCGCACACCGCGTGCATCAGGGG 2520  
QY 2473 CAGTGGCCCTTCGAGCGTCTTCTACGCTTTCATGTGCGCACACCGCGTGCATCAGGGG 2532  
Db 2521 CAACTGCTACGTTCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
QY 2533 CAACTGCTACGTTCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2592  
Db 2581 CAGCTGCTCCTACGCGGACATGAGAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
QY 2593 CAGCTGCTCCTACGCGGACATGAGAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2652  
Db 2641 GCTCTGCTGCTTGGTGGATGATTCTTGTGTTGAGACCTCACCTCACCGCCGCAAAAC 2700  
QY 2653 GCTCTGCTGCTTGGTGGATGATTCTTGTGTTGAGACCTCACCTCACCGCCGCAAAAC 2712  
Db 2701 CTTCTCAGAGACCCCTGGTCCGAGGTGCTCCTGAGTATGGCTCGTGGTGAATTCGGGAA 2760  
QY 2713 CTTCTCAGAGACCCCTGGTCCGAGGTGCTCCTGAGTATGGCTCGTGGTGAATTCGGGAA 2772  
Db 2761 GACAGTGTGAATTCCTCTGTAGAACAGAGCCCTGGGTGGACGCTTTTGTTCAGAT 2820  
QY 2773 GACAGTGTGAATTCCTCTGTAGAACAGAGCCCTGGGTGGACGCTTTTGTTCAGAT 2832  
Db 2821 GCGGCGCCACGGCTATTCCTCGTGGCTGCTGCTGGATACCGGACCCCTGGAGT 2880  
QY 2833 GCGGCGCCACGGCTATTCCTCGTGGCTGCTGCTGGATACCGGACCCCTGGAGT 2892  
Db 2881 GAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCGGCTCACCTTCAACCG 2940  
QY 2893 GAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCGGCTCACCTTCAACCG 2952  
Db 2941 CGCTTCAAGGCTGGGAGAACATCGCTCGCAAACTCTTTTGGGCTCTTCGGCTGAAGT 3000  
QY 2953 CGCTTCAAGGCTGGGAGAACATCGCTCGCAAACTCTTTTGGGCTCTTCGGCTGAAGT 3012  
Db 3001 TCACAGCTCTTCTGGATTTCAGGTGAACAGCCCTCCAGAGGGTGTGCACCAACATCTA 3060  
QY 3013 TCACAGCTCTTCTGGATTTCAGGTGAACAGCCCTCCAGAGGGTGTGCACCAACATCTA 3072  
Db 3061 CAAATCTCTCTGCTGAGCGGTACAGGTTTTCAGCATGTGTGCTGCACTCCCAATTTCA 3120  
QY 3073 CAAATCTCTCTGCTGAGCGGTACAGGTTTTCAGCATGTGTGCTGCACTCCCAATTTCA 3132  
Db 3121 TCAGCAAGTTTGGAGAACCCACATTTTCTCGCGCTCATCTCTGACACGCGCTCCCT 3180  
QY 3133 TCAGCAAGTTTGGAGAACCCACATTTTCTCGCGCTCATCTCTGACACGCGCTCCCT 3192  
Db 3181 CTGCTACTCCATCTTGAAGCCAAAGCAGGGATGTGCTGGGGGCCAAGGGCGCGC 3240  
QY 3193 CTGCTACTCCATCTTGAAGCCAAAGCAGGGATGTGCTGGGGGCCAAGGGCGCGC 3252  
Db 3241 CCGCCCTCTGCGCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGT 3300  
QY 3253 CCGCCCTCTGCGCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGT 3312  
Db 3301 GACTGACACCGTGTCACTACGTGCCACTCCTGGGTCACTCAGGACAGCCAGACGA 3360  
QY 3313 GACTGACACCGTGTCACTACGTGCCACTCCTGGGTCACTCAGGACAGCCAGACGA 3372  
Db 3361 GCTGAGTCGGAAGCTCCCGGGACGAGCTGACTGCTCTGGAGGCGGACCAACCCGCG 3420  
QY 3373 GCTGAGTCGGAAGCTCCCGGGACGAGCTGACTGCCCTGGAGGCGGACCAACCCGCG 3432  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTTGAATGTGGCCACCCGCGCCACAGCCAGGCCGA 3480  
QY 3433 ACTGCCCTCAGACTTCAAGACCATCTTGAATGTGGCCACCCGCGCCACAGCCAGGCCGA 3492  
Db 3481 GAGCAGACACGAGCCCTGTACGCGGCTCTACGTCGCCAGGAGGAGGGCGGCC 3540  
QY 3493 GAGCAGACACGAGCCCTGTACGCGGCTCTACGTCGCCAGGAGGAGGGCGGCC 3552  
|||||



QY 912 CCACCTCTTTGGAGGTTGCGTCTCTGGCAGCGGCCACTCCACCCATCCGTGGGCGGCC 971  
Db 963 AGCACACGCGGGCCCCCATCCACATGCGGGCCACACAGCTCCCTGGGACACGCTTGTG 1022  
QY 972 AGCACACGCGGGCCCCCATCCACATGCGGGCCACACAGCTCCCTGGGACACGCTTGTG 1031  
Db 1023 CCCGGGTGACGCGGAGACCAAGCACTTCTCTACTCTCTAGGGGACAAAGAGCAGCTGC 1082  
QY 1032 CCCGGGTGACGCGGAGACCAAGCACTTCTCTACTCTCTAGGGGACAAAGAGCAGCTGC 1091  
Db 1083 GGCCTCTCTCTACTCTAGCTCTCTAGGCGCCAGCTGACTGGCGCTCGGAGGCTCGTGG 1142  
QY 1092 GGCCTCTCTCTACTCTAGCTCTCTAGGCGCCAGCTGACTGGCGCTCGGAGGCTCGTGG 1151  
Db 1143 AGACCATTTTCTGGGTTCAGGCGCTTGATGCCAGGAGTCCCCGCGAGTTGCCCGGCC 1202  
QY 1152 AGACCATTTTCTGGGTTCAGGCGCTTGATGCCAGGAGTCCCCGCGAGTTGCCCGGCC 1211  
Db 1203 TGCCCCAGCGCTACTGGCAATGCGGCCCTGTGTTCTTGGAGCTGCTTGGGAACACAGCGC 1262  
QY 1212 TGCCCCAGCGCTACTGGCAATGCGGCCCTGTGTTCTTGGAGCTGCTTGGGAACACAGCGC 1271  
Db 1263 AGTGCCTCTACGGGTGCTCTCAAGAGCGCACTGCCGCTCGAGCTGCGGTCAACCCAG 1322  
QY 1272 AGTGCCTCTACGGGTGCTCTCAAGAGCGCACTGCCGCTCGAGCTGCGGTCAACCCAG 1331  
Db 1323 CAGCGGTGTCTGTGCCCGGAGAGCCCAAGGGCTGTGTGGCGGCCCGCCGAGGAGG 1382  
QY 1332 CAGCGGTGTCTGTGCCCGGAGAGCCCAAGGGCTGTGTGGCGGCCCGCCGAGGAGG 1391  
Db 1383 ACACAGACCCCGCTGCTGTGAGCTGTCTCCGCGACAGCAGCCCTTGGCAGGTGT 1442  
QY 1392 ACACAGACCCCGCTGCTGTGAGCTGTCTCCGCGACAGCAGCCCTTGGCAGGTGT 1451  
Db 1443 ACGGCTTCTGCGGGCGCTGCTGCGCGGCTGGTGCCTCCAGGCTCTTGGGGCTCCAGGC 1502  
QY 1452 ACGGCTTCTGCGGGCGCTGCTGCGCGGCTGGTGCCTCCAGGCTCTTGGGGCTCCAGGC 1511  
Db 1503 ACAACGACCGCGCTTCTCAGGAACACCAAGAAGTTTCTATCTCCCTGGGGAGCATGCCA 1562  
QY 1512 ACAACGACCGCGCTTCTCAGGAACACCAAGAAGTTTCTATCTCCCTGGGGAGCATGCCA 1571  
Db 1563 AGCTCTCGCTCAGGAGCTGACGTGGAGATGACGTGCGGGGCTGCGCTTGGGTGGCGCA 1622  
QY 1572 AGCTCTCGCTCAGGAGCTGACGTGGAGATGACGTGCGGGGCTGCGCTTGGGTGGCGCA 1631  
Db 1623 GGAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGCAGCCGCTGCTGAGGAGATCCTGG 1682  
QY 1632 GGAGCCAGGGGTTGGCTGTGTTCCGGCCGACAGCAGCCGCTGCTGAGGAGATCCTGG 1691  
Db 1683 CCAAGTTCTCTGACTGGGTGATGAGTGTGACGTGCTGAGCTGCTCAGGTCTTTCTTTT 1742  
QY 1692 CCAAGTTCTCTGACTGGGTGATGAGTGTGACGTGCTGAGCTGCTCAGGTCTTTCTTT 1751  
Db 1743 ATGTCAGGAGACACAGTTTCAAAGAACACAGGCTCTTTTCTACCGAAGAGTGTCTGGA 1802  
QY 1752 ATGTCAGGAGACACAGTTTCAAAGAACACAGGCTCTTTTCTACCGAAGAGTGTCTGGA 1811  
Db 1803 GCAAGTTCAAAGGATGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGCTGT 1862  
QY 1812 GCAAGTTCAAAGGATGGAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGCTGT 1871  
Db 1863 CGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGGCGCGCTGCTGACGTCTCAGACTCC 1922  
QY 1872 CGGAAGCAGAGGTCAGGAGCATCGGGAAGCCAGGCGCGCTGCTGACGTCTCAGACTCC 1931  
Db 1923 GCTTCATCCCCAAGCTGACGGGTCGGGCCGATTTGAACATGACATGACCTCTGCTGGAG 1982  
QY 1932 GCTTCATCCCCAAGCTGACGGGTCGGGCCGATTTGAACATGACATGACCTCTGCTGGAG 1991  
Db 1983 CCAGAAGCTTCCGAGAGAAAAGGCGGAGCGTCTACCTCGAGGTGAAGCACTGT 2042  
QY 1992 CCAGAAGCTTCCGAGAGAAAAGGCGGAGCGTCTACCTCGAGGTGAAGCACTGT 2051

Db 2043 TCAGCGTGTCTAACTAGAGGGGCGCGCGCCCGGCTCTCTGGGCGCCTCTGTGTGCTGG 2102  
QY 2052 TCAGCGTGTCTAACTAGAGGGGCGCGCGCCCGGCTCTCTGGGCGCCTCTGTGTGCTGG 2111  
Db 2103 GCCTGGACGATATCCACAGGGCCTGGCGACCTTCGTGCTGCGTGTGCGGGCCAGGACC 2162  
QY 2112 GCCTGGACGATATCCACAGGGCCTGGCGACCTTCGTGCTGCGTGTGCGGGCCAGGACC 2171  
Db 2163 CGCGCGCTGAGCTGTACTTTTCAAGGTGATGTAGCGGGCGGTAGACACCATCCCC 2222  
QY 2172 CGCGCGCTGAGCTGTACTTTTCAAGGTGATGTAGCGGGCGGTAGACACCATCCCC 2231  
Db 2223 AGACAGGCTACGAGGTCATCGCAGCATCATCAAAACCCAGAACACGCTACTGCGTC 2282  
QY 2232 AGACAGGCTACGAGGTCATCGCAGCATCATCAAAACCCAGAACACGCTACTGCGTC 2291  
Db 2283 GTCGGTATGCCGTGGTCCAGAGGCCCGCCATGGGACGTCGCGAGGCCCTTCAAGACC 2342  
QY 2292 GTCGGTATGCCGTGGTCCAGAGGCCCGCCATGGGACGTCGCGAGGCCCTTCAAGACC 2351  
Db 2343 ACGTCTCTACTTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACTGCGAG 2402  
QY 2352 ACGTCTCTACTTTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACTGCGAG 2411  
Db 2403 AGACAGCCCGCTGAGGGATGCGGTCTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCA 2462  
QY 2412 AGACAGCCCGCTGAGGGATGCGGTCTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCA 2471  
Db 2463 GCAGTGGCCTTTCGAGCTTCTTCTAGCGTTCTATGTGCCACAGCGCGTGGCATCAGGG 2522  
QY 2472 GCAGTGGCCTTTCGAGCTTCTTCTAGCGTTCTATGTGCCACAGCGCGTGGCATCAGGG 2531  
Db 2523 GCAAGTCTACGTCAGTCGAGGGATCCCGAGGGCTCCATCCTCTCCACGCTGCTCT 2582  
QY 2532 GCAAGTCTACGTCAGTCGAGGGATCCCGAGGGCTCCATCCTCTCCACGCTGCTCT 2591  
Db 2583 GCAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGCGGGGACGGGC 2642  
QY 2592 GCAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGCGGGGACGGGC 2651  
Db 2643 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTGACACCTCACTCAACCCAGCGGAAAA 2702  
QY 2652 TGCTCTGCGCTTGGTGGATGATTTCTTGTGTGACACCTCACTCAACCCAGCGGAAAA 2711  
Db 2703 CCTTCTCAGAGCCCTGTCGCGAGGTGTCCTGAGTATGCTGCTGCTGTAACCTTGGGA 2762  
QY 2712 CCTTCTCAGAGCCCTGTCGCGAGGTGTCCTGAGTATGCTGCTGCTGCTGTAACCTTGGGA 2771  
Db 2763 AGACAGTGTGAACTTCCCTGTAGAAGACAGGCGCCCTGGGTGGCACGCGCTTTTGTTCAGA 2822  
QY 2772 AGACAGTGTGAACTTCCCTGTAGAAGACAGGCGCCCTGGGTGGCACGCGCTTTTGTTCAGA 2831  
Db 2823 TGCGGGCCACAGGCTATTCCTCTGTCGCGCTGCTGCTGCTGATACCCGAGCCCTGAGG 2882  
QY 2832 TGCGGGCCACAGGCTATTCCTCTGTCGCGCTGCTGCTGCTGATACCCGAGCCCTGAGG 2891  
Db 2883 TGACAGCGACTACTCCAGCTATGCCGGACCTCCATCAGACGACGCTCACTCACTCAAC 2942  
QY 2892 TGACAGCGACTACTCCAGCTATGCCGGACCTCCATCAGACGACGCTCACTCACTCAAC 2951  
Db 2943 CGCGCTTCAAAGCTGGGAGGAACATGCTCGCAAACTCTTTTGGGGTCTTTCGGGTGAAGT 3002  
QY 2952 CGCGCTTCAAAGCTGGGAGGAACATGCTCGCAAACTCTTTTGGGGTCTTTCGGGTGAAGT 3011  
Db 3003 GTCACAGCCTTCTGCTGATTTGAGGTGAACAGCCTCCAGAGGCTGACACCAATCT 3062  
QY 3012 GTCACAGCCTTCTGCTGATTTGAGGTGAACAGCCTCCAGAGGCTGACACCAATCT 3071  
Db 3063 ACAAGATCTCTGCTGACGCGTACAGGTTTCAAGCATGTGCTGACGCTCCCATTTTC 3122  
QY 3072 ACAAGATCTCTGCTGACGCGGTACAGGTTTCAAGCATGTGCTGACGCTCCCATTTTC 3131







2532 GCAAGTCTTACGTCCAGTGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCT 2591  
2575 GCAGCCTGTCTACGCGCATATGGAAACAAGCTGTTTGGGGGATTCGCGGGACGGGC 2634  
2592 GCAGCCTGTCTACGCGCATATGGAAACAAGCTGTTTGGGGGATTCGCGGGACGGGC 2651  
2635 TGCTCTGGTGTGGTGGATGATTTCTTGTGTGACACTCACCTACCCAGCGCAAAA 2694  
2652 TGCTCTGGTGTGGTGGATGATTTCTTGTGTGACACTCACCTACCCAGCGCAAAA 2711  
2695 CTTCTCTAGGACCTGGTCCGAGGTGTCCTGAGTATGCTGCGTGGTGAAGTTCGGGA 2754  
2712 CTTCTCTAGGACCTGGTCCGAGGTGTCCTGAGTATGCTGCGTGGTGAAGTTCGGGA 2771  
2755 AGACAGTGTGAATTCCTCTGTAGAAGAGAGCCCTGGGTGGCAGGGCTTTTGTTCAGA 2814  
2772 AGACAGTGTGAATTCCTCTGTAGAAGAGAGCCCTGGGTGGCAGGGCTTTTGTTCAGA 2831  
2815 TGCCGGCCACAGCCCTATTCCCTGTGGTGGGCTGCTGTGTGATACCCGGACCTGGAGG 2874  
2832 TGCCGGCCACAGCCCTATTCCCTGTGGTGGGCTGCTGTGTGATACCCGGACCTGGAGG 2891  
2875 TGACAGAGGACTACTCCAGTATGCCGAGCCTCCATCAGAGCCAGTCTCACCTTCAACC 2934  
2892 TGACAGAGGACTACTCCAGTATGCCGAGCCTCCATCAGAGCCAGTCTCACCTTCAACC 2951  
2935 GCGGCTTCAAGGCTGGAGGAACATCGCTGCAAACTCTTTGGGGTCTTGGCGCTGAAGT 2994  
2952 GCGGCTTCAAGGCTGGAGGAACATCGCTGCAAACTCTTTGGGGTCTTGGCGCTGAAGT 3011  
2995 GTACAGCCTGTTTCTGGATTTCCAGGTGAACAGCCTCCAGAGCGGTGTGACCAACATCT 3054  
3012 GTACAGCCTGTTTCTGGATTTCCAGGTGAACAGCCTCCAGAGCGGTGTGACCAACATCT 3071  
3055 ACAAGATCCTGCTGCGAGGCTACAGTTTCAAGCATGTGCTGCAGCTCCCATTTTC 3114  
3072 ACAAGATCCTGCTGCGAGGCTACAGTTTCAAGCATGTGCTGCAGCTCCCATTTTC 3131  
3115 ATCAGCAAGTTTGGAAAGACCCACACATTTTCTGGCGGTCTCTGTGACAGCGCTCC 3174  
3132 ATCAGCAAGTTTGGAAAGACCCACACATTTTCTGGCGGTCTCTGTGACAGCGCTCC 3191  
3175 TGTGCTACTCCATCTCTGAAAGCAAGAGCAGGATGTGCTGGGGGCAAGGGCGCG 3234  
3192 TGTGCTACTCCATCTCTGAAAGCAAGAGCAGGATGTGCTGGGGGCAAGGGCGCG 3251  
3235 CCGGCGCTGTGCGCTCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGC 3294  
3252 CCGGCGCTGTGCGCTCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGC 3311  
3295 TGACTCGACACCGTGTACCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3354  
3312 TGACTCGACACCGTGTACCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3371  
3355 AGCTGAGTGGAAAGTCTCCGGGGACAGCGTACTGCTGCTGAGGGCGGCAAGCCCGG 3414  
3372 AGCTGAGTGGAAAGTCTCCGGGGACAGCGTACTGCTGCTGAGGGCGGCAAGCCCGG 3431  
3415 CACTGCGCTCAGACTTCAAGACCATCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3474  
3432 CACTGCGCTCAGACTTCAAGACCATCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3491  
3475 AGAGCAGACACAGCAGCCCTGTGACGCGGCTCTACGTCCAGGGAGGGAGGGCGGC 3534  
3492 AGAGCAGACACAGCAGCCCTGTGACGCGGCTCTACGTCCAGGGAGGGAGGGCGGC 3551  
3535 CCACACCCAGGCGCCAGCCGCTGGAGTCTGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3594  
3552 CCACACCCAGGCGCCAGCCGCTGGAGTCTGAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3611  
3595 GCATGTCCGGCTCAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3654

3612 GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3671  
3655 TGAGTGTCCAGCACACCTGCGTCTTCACTTCCACAGGCTGGGCTCGGCTCCACCC 3714  
3672 TGAGTGTCCAGCACACCTGCGTCTTCACTTCCACAGGCTGGGCTCGGCTCCACCC 3731  
3715 AGGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCATAGGAATAGTCCATC 3774  
3732 AGGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCATAGGAATAGTCCATC 3791  
3775 CCAGAT 3781  
3792 CCCTGAT 3798

RESULT 8  
ID US-08-851-843-173 STANDARD; DNA; UNC; 4029 BP.  
AC xxxxxx  
DT Sequence 173, Application US/08851843  
CC Sequence 173, Application US/08851843  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Novel Telomerase  
CC NUMBER OF SEQUENCES: 223  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 173:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 4029 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cdna





QY 1992 CCAGAACGTTCCGACAGAAAAAGAGCGCGGACGCTCTCACCTCGAGGGTGAAGGCACTGT 2051  
Db 2035 TCAGGCTGCTCAACTACGAGCGGCGCGCGCCGCTCTCTGGCGCCTCTGTGCTGG 2094  
QY 2052 TCAGGCTGCTCAACTACGAGCGGCGCGCGCCGCTCTCTGGCGCCTCTGTGCTGG 2111  
Db 2095 GCCTGGACGATATCCAGAGGCTTGGCGCACCTTCGTGCTGGGTGTGGGGCCAGGACC 2154  
QY 2112 GCCTGGACGATATCCAGAGGCTTGGCGCACCTTCGTGCTGGGTGTGGGGCCAGGACC 2171  
Db 2155 CGCCGCTTGAGCTGTACTTTGTCAAGGTGGATGTACGCGGCGCTACGACACCATCCGCC 2214  
QY 2172 CGCCGCTTGAGCTGTACTTTGTCAAGGTGGATGTACGCGGCGCTACGACACCATCCGCC 2231  
Db 2215 AGGACAGGCTCACGAGGTCTATCGCCAGCATCATCAAAACCCAGAACACGCTACTCGTGC 2274  
QY 2232 AGGACAGGCTCACGAGGTCTATCGCCAGCATCATCAAAACCCAGAACACGCTACTCGTGC 2291  
Db 2275 GTCGATATGCGGTGCTCCAGAGGCGCGCCATGGCGACGTCGCGAAGGCTTCAAGAGCC 2334  
QY 2292 GTCGATATGCGGTGCTCCAGAGGCGCGCCATGGCGACGTCGCGAAGGCTTCAAGAGCC 2351  
Db 2335 AGCTCTCTACCTTTGACAGACCTCCAGCGGTACATCGGACGTTGCTGCTCACCTGCAGG 2394  
QY 2352 AGCTCTCTACCTTTGACAGACCTCCAGCGGTACATCGGACGTTGCTGCTCACCTGCAGG 2411  
Db 2395 ANAACAGCCCGGTGAGGATGCCGTGTCATCGAGCAGAGCTCTCCCTGATGAGGCCA 2454  
QY 2412 AGAACAGCCCGGTGAGGATGCCGTGTCATCGAGCAGAGCTCTCCCTGATGAGGCCA 2471  
Db 2455 GCAGTGGCCTCTTCAAGCTCTTCCACGTTTACGTTTATGTCACACACGCGTGGCGATCAGG 2514  
QY 2472 GCAGTGGCCTCTTCAAGCTCTTCCACGTTTATGTCACACACGCGTGGCGATCAGG 2531  
Db 2515 GCAAGTCTCTACGTCAGTCCAGAGGATCCCGACGGGTCCATCTCTCCAGGCTGCTCT 2574  
QY 2532 GCAAGTCTCTACGTCAGTCCAGAGGATCCCGACGGGTCCATCTCTCCAGGCTGCTCT 2591  
Db 2575 GCAGCTGTGCTACGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGGACGGGC 2634  
QY 2592 GCAGCTGTGCTACGCGACATGAGAACAAAGCTGTTTGGGGGATTCGGCGGACGGGC 2651  
Db 2635 TGCTCTGGTTTGGTGAATTTCTTGTGTTGACACCTCACCTCACCCACGCGAATA 2694  
QY 2652 TGCTCTGGTTTGGTGAATTTCTTGTGTTGACACCTCACCTCACCCACGCGAATA 2711  
Db 2695 CTTCTCTCAGGACCTGTGTCAGAGTTCCTTGATGATGGCTGCTGATGTTGCGGA 2754  
QY 2712 CTTCTCTCAGGACCTGTGTCAGAGTTCCTTGATGATGGCTGCTGATGTTGCGGA 2771  
Db 2755 AGACAGTGTGAATTTCCCTGTAGAACAGAGGCGCTGGGTGGCACGGCTTTTGTTCAGA 2814  
QY 2772 AGACAGTGTGAATTTCCCTGTAGAACAGAGGCGCTGGGTGGCACGGCTTTTGTTCAGA 2831  
Db 2815 TGCCGGCCACGGCTATTTCCTCTGGTGGGCGCTGCTGCTGATACCCGACCTTGAGG 2874  
QY 2832 TGCCGGCCACGGCTATTTCCTCTGGTGGGCGCTGCTGCTGATACCCGACCTTGAGG 2891  
Db 2875 TGACAGGCACTACTCCAGTATGCCGACCTCCATCAGAGCCAGTCTCACCTTCAACC 2934  
QY 2892 TGACAGGCACTACTCCAGTATGCCGACCTCCATCAGAGCCAGTCTCACCTTCAACC 2951  
Db 2935 GGGGCTTCAAGGCTGGAGAACATGCGTCGCAAACTCTTTGGGCTTTCGGGCTGAAGT 2994  
QY 2952 GGGGCTTCAAGGCTGGAGAACATGCGTCGCAAACTCTTTGGGCTTTCGGGCTGAAGT 3011  
Db 2995 GTACAGGCTGTTTCTGGATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT 3054  
QY 3012 GTACAGGCTGTTTCTGGATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT 3071  
Db 3055 ACAAGATCTCTGCTGACGGGCTACAGGTTTACGATGTTGCTGAGCTTCCATTTTC 3114  
QY 3072 ACAAGATCTCTGCTGACGGGCTACAGGTTTACGATGTTGCTGAGCTTCCATTTTC 3131

Db 3115 ATCAGCAAGTTTGAAGAAACCCACATTTTCTCGCGGTCACTCTGTACACGCGCTGCC 3174  
QY 3132 ATCAGCAAGTTTGAAGAAACCCACATTTTCTCGCGGTCACTCTGTACACGCGCTGCC 3191  
Db 3175 TCTGCTACTCCTCTGAAAGCCAAAGAACGACGAGGATGCTGCTGGGGCCAAAGGCGCG 3234  
QY 3192 TCTGCTACTCCTCTGAAAGCCAAAGAACGACGAGGATGCTGCTGGGGCCAAAGGCGCG 3251  
Db 3235 CGGGCCCTCTGCCCTCCGAGGCGGTGACGTGGCTGTGCACCAAGACATTTCTGTCTCAAGC 3294  
QY 3252 CGGGCCCTCTGCCCTCCGAGGCGGTGACGTGGCTGTGCACCAAGACATTTCTGTCTCAAGC 3311  
Db 3295 TGACTCGACACCGTGTCACTTACCTTACCTTCTTGGGTCTACTCAGGACACCCAGACGC 3354  
QY 3312 TGACTCGACACCGTGTCACTTACCTTACCTTCTTGGGTCTACTCAGGACACCCAGACGC 3371  
Db 3355 AGCTGAGTCGGAAGCTCCCGGGGAGGAGCTGACTGCCCTGGAGGCCGACGCAACCGCG 3414  
QY 3372 AGCTGAGTCGGAAGCTCCCGGGGAGGAGCTGACTGCCCTGGAGGCCGACGCAACCGCG 3431  
Db 3415 CACTGCCCTCAGACTTCAAGACCTTCTTGGACTGATGSCCACCCTGATGSCCACCAGCGCGC 3474  
QY 3432 CACTGCCCTCAGACTTCAAGACCTTCTTGGACTGATGSCCACCCTGATGSCCACCAGCGCGC 3491  
Db 3475 AGAGCAGACACCCAGACGCGCTGTCAAGCGGCTTCTACGTCACAGGAGGAGGCGCGC 3534  
QY 3492 AGAGCAGACACCCAGACGCGCTGTCAAGCGGCTTCTACGTCACAGGAGGAGGCGCGC 3551  
Db 3535 CCACACCCAGCGCGCAGCGCTGGAGTCTGAGGCTGAGTGTGTTGGCGAGGCT 3594  
QY 3552 CCACACCCAGCGCGCAGCGCTGGAGTCTGAGGCTGAGTGTGTTGGCGAGGCT 3611  
Db 3595 GCATGTCGGCTGAAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3654  
QY 3612 GCATGTCGGCTGAAGGCTGAGTGTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3671  
Db 3655 TGAGTGTCCAGCACACCTTCCGCTTCTACTTCCACAGGCTGGCGCTGCGCTTCCACGCC 3714  
QY 3672 TGAGTGTCCAGCACACCTTCCGCTTCTACTTCCACAGGCTGGCGCTGCGCTTCCACGCC 3731  
Db 3715 AGGGCAGCTTTTCTTACACAGGCGCGGCTTCCACTTCCACATAGGAATAGTCTCATC 3774  
QY 3732 AGGGCAGCTTTTCTTACACAGGCGCGGCTTCCACTTCCACATAGGAATAGTCTCATC 3791  
Db 3775 CCCAGAT 3781  
QY 3792 CCTGAT 3798

## RESULT 9

ID US-08-912-951-4 STANDARD; DNA; UNC; 3855 BP.

AC xxxxx

DE Sequence 4, Application US/08912951

CC Sequence 4, Application US/08912951

CC GENERAL INFORMATION:

CC APPLICANT: Cech, Thomas R.

CC APPLICANT: Lingner, Joachim

CC APPLICANT: Nakamura, Toru

CC APPLICANT: Chapman, Karen B.

CC APPLICANT: Morin, Gregg B.

CC APPLICANT: Harley, Calvin

CC APPLICANT: Andrews, William H.

CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

CC NUMBER OF SEQUENCES: 335

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, 8th Floor

CC CITY: San Francisco

CC STATE: California

CC COUNTRY: United States of America







QY 613 TCAGGCCGCGCCGACACAGCTAGTGGACCCGGAAGCGCTCTGGGATGCGAAGCGGC 672  
Db 661 CTGGAACCATAGGCTAGGAGGCGGGGTCCCTCTGGGCTTCCAGCCCGGGTGGAG 720  
QY 673 CTGGAACCATAGGCTAGGAGGCGGGGTCCCTCTGGGCTTCCAGCCCGGGTGGAG 732  
Db 721 GAGCGCGGGGAGTCCAGCGAAGCTCTGCGGTTCGCCCAAGAGGCCAGGCTGSCGC 780  
QY 733 GAGCGCGGGGAGTCCAGCGAAGCTCTGCGGTTCGCCCAAGAGGCCAGGCTGSCGC 792  
Db 781 TGCCCTTAGCCGAGCGGAGCGCCCGTTCGGCAGGGTCTCTGGGCCACCCCGGSCAGGAC 840  
QY 793 TGCCCTTAGCCGAGCGGAGCGCCCGTTCGGCAGGGTCTCTGGGCCACCCCGGSCAGGAC 852  
Db 841 GCGTGACCGAGTACCGTGGTTCCTGTGTGTGTACCTGCGCAGACCCCGCAAGAAGC 900  
QY 853 GCGTGACCGAGTACCGTGGTTCCTGTGTGTGTACCTGCGCAGACCCCGCAAGAAGC 912  
Db 901 CACCTCTTTGAGGGTGCCTCTCTGACGCGCCACTCCACCCATCCGCTGGGCGGCCA 960  
QY 913 CACCTCTTTGAGGGTGCCTCTCTGACGCGCCACTCCACCCATCCGCTGGGCGGCCA 972  
Db 961 GCACACGCGGGCCCCCATCCACATCGCGGCCACCACTCCCTGGACACGCTTGTCC 1020  
QY 973 GCACACGCGGGCCCCCATCCACATCGCGGCCACCACTCCCTGGACACGCTTGTCC 1032  
Db 1021 CCGGTTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAAGAGAGCTGCG 1080  
QY 1033 CCGGTTAGCGCGAGACCAAGCACTTCTCTACTCTCAGGGGACAAAGAGAGCTGCG 1092  
Db 1081 GCCTCTCTCTACTAGCTCTCTGAGGCCAGCTTCTGAGGCTCGGAGGCTCTGTGA 1140  
QY 1093 GCCTCTCTCTACTAGCTCTCTGAGGCCAGCTTCTGAGGCTCGGAGGCTCTGTGA 1152  
Db 1141 GACCATCTTTCTGGTTCAGGGCCCTGGATGCCAGGAGCTCCCGCAGGTTGCCCGCCT 1200  
QY 1153 GACCATCTTTCTGGTTCAGGGCCCTGGATGCCAGGAGCTCCCGCAGGTTGCCCGCCT 1212  
Db 1201 GCGCCAGCGTACTGGCAATAGGCGCCCTGTTTCTGGAGCTGCTGGGAACCAACGCGCA 1260  
QY 1213 GCGCCAGCGTACTGGCAATAGGCGCCCTGTTTCTGGAGCTGCTGGGAACCAACGCGCA 1272  
Db 1261 GTGCCCTAGCGGTGCTCTCAGAGCGCACTGCCCTCGGAGCTCGGCTCAGCCGAGC 1320  
QY 1273 GTGCCCTAGCGGTGCTCTCAGAGCGCACTGCCCTCGGAGCTCGGCTCAGCCGAGC 1332  
Db 1321 AGCGGTGCTCTGTCCCGGAGAACGCCAGGGCTCTGTGGCGGCCCGCAGGAGGAGGA 1380  
QY 1333 AGCGGTGCTCTGTCCCGGAGAACGCCAGGGCTCTGTGGCGGCCCGCAGGAGGAGGA 1392  
Db 1381 CACAGACCCCGCTGCTGT 1440  
QY 1393 CACAGACCCCGCTGCTGT 1452  
Db 1441 CGGCTTGTGGGCGCTGCTGTGGCGGGCTGTGGCGGCCAGAGGCTCTGGGCTTCAGGCA 1500  
QY 1453 CGGCTTGTGGGCGCTGCTGTGGCGGGCTGTGGCGGCCAGAGGCTCTGGGCTTCAGGCA 1512  
Db 1501 CAACGAACCGCGCTTCTCTAGGAACACCAAGAGTTCTCTCCTTGGGGAAGCATGCCAA 1560  
QY 1513 CAACGAACCGCGCTTCTCTAGGAACACCAAGAGTTCTCTCCTTGGGGAAGCATGCCAA 1572  
Db 1561 GCTCTGCTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1620  
QY 1573 GCTCTGCTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1632  
Db 1621 GAGCCAGGGTGGCTGTGTTCGGGCGGAGAGCACCGCTCTGGGTGAGGAGTCTCTGGC 1680  
QY 1633 GAGCCAGGGTGGCTGTGTTCGGGCGGAGAGCACCGCTCTGGGTGAGGAGTCTCTGGC 1692  
Db 1681 CAAGTTCCTGCACTGGCTGATGATGTGTACGTCGTGAGCTGCTCAGGCTTCTCTTTTA 1740  
QY 1693 CAAGTTCCTGCACTGGCTGATGATGTGTACGTCGTGAGCTGCTCAGGCTTCTCTTTTA 1752

Db 1741 TGTACGGAGACACACGCTTTCAAAGAACAGGCTCTTTTCTTACCAGGAGTCTCTGGAG 1800  
QY 1753 TGTACGGAGACACACGCTTTCAAAGAACAGGCTCTTTTCTTACCAGGAGTCTCTGGAG 1812  
Db 1801 CAAGTTGCAAGCATTTGGAATCAGACAGCACTTTGAAGAGGGTGCAGCTGGGAGCTGTC 1860  
QY 1813 CAAGTTGCAAGCATTTGGAATCAGACAGCACTTTGAAGAGGGTGCAGCTGGGAGCTGTC 1872  
Db 1861 GGAAGCAGAGGTTCAGACAGCATCGGGAAGCCAGGCCCGCCCTCTGACGTCTCAGACTCCG 1920  
QY 1873 GGAAGCAGAGGTTCAGACAGCATCGGGAAGCCAGGCCCGCCCTCTGACGTCTCAGACTCCG 1932  
Db 1921 CTTATCTCCCAAGCCCTGAGGGCTGGGCGGATTTGTAACATGGACTACGTCCTGGGAGC 1980  
QY 1933 CTTATCTCCCAAGCCCTGAGGGCTGGGCGGATTTGTAACATGGACTACGTCCTGGGAGC 1992  
Db 1981 CAGAAGCTTCCCGCAGAGAAAGAGGCGCGGCGCTCTGAGGGGCTCTGCTGTGCTGG 2040  
QY 1993 CAGAAGCTTCCCGCAGAGAAAGAGGCGCGGCGCTCTGAGGGGCTCTGCTGTGCTGG 2052  
Db 2041 CAGCGTGTCTCAACTAGCAGCGGCGGCGGCGGCGCTCTGAGGGGCTCTGCTGTGCTGG 2100  
QY 2053 CAGCGTGTCTCAACTAGCAGCGGCGGCGGCGGCGCTCTGAGGGGCTCTGCTGTGCTGG 2112  
Db 2101 CTTGGAGATATCCACAGGCGCTGGGCGACCTTCTGTGCTGCTGTGCGGCGCCAGAGACC 2160  
QY 2113 CTTGGAGATATCCACAGGCGCTGGGCGACCTTCTGTGCTGCTGTGCGGCGCCAGAGACC 2172  
Db 2161 GCGCGCTGAGCTGACTTGTCAAGTGGATGTGAGGGGCGCTGAGACACCATTCGCCCA 2220  
QY 2173 GCGCGCTGAGCTGACTTGTCAAGTGGATGTGAGGGGCGCTGAGACACCATTCGCCCA 2232  
Db 2221 GGACAGGCTCACGAGGCTCATGCCAGCATCATCAAAACCCAGAACACGCTACTGCTGCG 2280  
QY 2233 GGACAGGCTCACGAGGCTCATGCCAGCATCATCAAAACCCAGAACACGCTACTGCTGCG 2292  
Db 2281 TCGGTATGCGGTGGTCCAGAGGCCCGCATGGGCGACGTCGCCAAGCCCTTCAAGAGCCA 2340  
QY 2293 TCGGTATGCGGTGGTCCAGAGGCCCGCATGGGCGACGTCGCCAAGCCCTTCAAGAGCCA 2352  
Db 2341 CGTC----- 2344  
QY 2353 CGTCTTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAGGA 2412  
Db 2344 ----- 2344  
QY 2413 GACCAGCCGCTCAGGGATGCCCTCGTCATCGACAGAGAGCTCCTCCTGATGAGGCCAG 2472  
Db 2344 ----- 2344  
QY 2473 CAGTGGCCTCTTCGACGCTTCTTACGCTTTCATGTGCCACACGCGCTGGGCATCAGGGG 2532  
Db 2345 ----- 2345  
QY 2533 CAAAGTCTTAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2592  
Db 2399 CAGCCTGTGCTACGGGACATGAGAACAGCTTTTGGGGGATTCGGGGGACGCGGCT 2458  
QY 2593 CAGCCTGTGCTACGGGACATGAGAACAGCTTTTGGGGGATTCGGGGGACGCGGCT 2652  
Db 2459 GCTCCTGCGTTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2518  
QY 2653 GCTCCTGCGTTTGGTGGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2712  
Db 2519 CTTCTCAGAGACCCCTGGTCCGAGGTGTCCTGTAGTATGGCTGGTGGTGAACCTTGGGAA 2578  
QY 2713 CTTCTCAGAGACCCCTGGTCCGAGGTGTCCTGTAGTATGGCTGGTGGTGAACCTTGGGAA 2772  
Db 2579 GACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTCTTTCAGAT 2638  
QY 2773 GACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTCTTTCAGAT 2832

Db 2639 GCGGGCCACGGCTATTCCCTGGTGGCGCTGCTGATACCGGACCTGGAGGT 2698  
QY 2833 GCGGGCCACGGCTATTCCCTGGTGGCGCTGCTGATACCGGACCTGGAGGT 2892  
Db 2699 GCAGAGGACTACTCCAGCTATGCCCCAGACTCCATCAGAGCCAGTCTCACCTTCACCG 2758  
QY 2893 GCAGAGGACTACTCCAGCTATGCCCCAGACTCCATCAGAGCCAGTCTCACCTTCACCG 2952  
Db 2759 CGGCTTCAAGGCTGGGAGAACATGCTCGCAAACTCTTTGGGCTCTTGGGCTGAAGTG 2818  
QY 2953 CGGCTTCAAGGCTGGGAGAACATGCTCGCAAACTCTTTGGGCTCTTGGGCTGAAGTG 3012  
Db 2819 TCACAGCTCTGTTCTGATTTGAGGTGAACAGCTCCAGACGCTGACCAACATCTA 2878  
QY 3013 TCACAGCTCTGTTCTGATTTGAGGTGAACAGCTCCAGACGCTGACCAACATCTA 3072  
Db 2879 CAAGATCCTCTGCTGAGGCTACAGGTTTCAGCATGTGTGTCAGCTCCCATTTCA 2938  
QY 3073 CAAGATCCTCTGCTGAGGCTACAGGTTTCAGCATGTGTGTCAGCTCCCATTTCA 3132  
Db 2939 TCAGCAAGTTTGAAGAACCCACATTTTCTCGCGGTCACTCTGACACGGCTCCCT 2998  
QY 3133 TCAGCAAGTTTGAAGAACCCACATTTTCTCGCGGTCACTCTGACACGGCTCCCT 3192  
Db 2999 CTGCTACTCCTGAAAGCCAGGACGAGGATGCTGCGGGCCAGGCGCGC 3058  
QY 3193 CTGCTACTCCTGAAAGCCAGGACGAGGATGCTGCGGGCCAGGCGCGC 3252  
Db 3059 CGGCTCTCTGCTGAGGCTGCAAGTGGCTGTGCAACAGCATTTCTGCTCAAGCT 3118  
QY 3253 CGGCTCTCTGCTGAGGCTGCAAGTGGCTGTGCAACAGCATTTCTGCTCAAGCT 3312  
Db 3119 GACTCGACCGTGTACCTAGTGGCACTCTGCGGTGCTGAGGACGCGCAGAGCA 3178  
QY 3313 GACTCGACCGTGTACCTAGTGGCACTCTGCGGTGCTGAGGACGCGCAGAGCA 3372  
Db 3179 GCTGAGTGGAAAGTCCCGGGAGCGCTGACTGCTGGAGCGCGCAGCCACCGGC 3238  
QY 3373 GCTGAGTGGAAAGTCCCGGGAGCGCTGACTGCTGGAGCGCGCAGCCACCGGC 3432  
Db 3239 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGCGCCACCGCCAGAGCGCA 3298  
QY 3433 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGTGATGCGCCACCGCCAGAGCGCA 3492  
Db 3299 GAGGACACACAGAGCCCTGTACGCGCGGCTTACGTCCCGAGGAGGAGGCGGCC 3358  
QY 3493 GAGGACACACAGAGCCCTGTACGCGCGGCTTACGTCCCGAGGAGGAGGCGGCC 3552  
Db 3359 CACACCCAGGCGCGCAGCTGGAGTCTGAGGCTGAGTGTGTTGGCGGAGCGCTG 3418  
QY 3553 CACACCCAGGCGCGCAGCTGGAGTCTGAGGCTGAGTGTGTTGGCGGAGCGCTG 3512  
Db 3419 CATGTCCGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3478  
QY 3613 CATGTCCGGCTGAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3672  
Db 3479 GAGTGTCCAGACACCTGCGCTTCTACTTCCCGACAGGCTGGCGCTGCGCTCCACCCCA 3538  
QY 3673 GAGTGTCCAGACACCTGCGCTTCTACTTCCCGACAGGCTGGCGCTGCGCTCCACCCCA 3732  
Db 3539 GGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCATAGGATAGTCCATCC 3598  
QY 3733 GGCGAGCTTTTCTCACCAGGAGCGCGCTTCCACTCCCATAGGATAGTCCATCC 3792  
Db 3599 CCAGAT 3604  
QY 3793 CCTGAT 3798

RESULT 11  
ID US-08-951-733-13 STANDARD; DNA; UNC; 2848 BP.  
AC xxxxxx  
DT

DE Sequence 13, Application US/08951733  
CC Sequence 13, Application US/08951733  
CC GENERAL INFORMATION:  
CC APPLICANT: Harrington, Lea A.  
CC APPLICANT: Robinson, Murray O.  
CC TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS  
CC NUMBER OF SEQUENCES: 44  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Amgen Inc.  
CC STREET: One Amgen Center Drive  
CC CITY: Thousand Oaks  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 91320-1789  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/951,733  
CC FILING DATE: 16-OCT-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/873,039  
CC FILING DATE: 11-JUN-1997  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/751,189  
CC FILING DATE: 15-NOV-1996  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Oleski, Nancy A.  
CC REGISTRATION NUMBER: 34,688  
CC REFERENCE/DOCKET NUMBER: A-433B  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (805) 447-6504  
CC TELEFAX: (805) 499-8011  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2848 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC SEQUENCE 2848 BP; 437 A; 978 C; 945 G; 488 T; 0 OTHER.

Query Match 75.0%; Score 2848; DB 22; Length 2848;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 2848; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CAGCGCTCGGGCAGCGCTGCTGCTGCGACGCTGGAGACCTGCGCGCCGACCC 60  
QY 2 CAGCGCTCGGGCAGCGCTGCTGCTGCGACGCTGGAGACCTGCGCGCCGACCC 61  
Db 61 CCGCGCATCGCGCGCTCCCGCTGCCAGCGCTGCTGCTGCTGCGCAGCACTAC 120  
QY 62 CCGCGCATCGCGCGCTCCCGCTGCCAGCGCTGCTGCTGCTGCGCAGCACTAC 121  
Db 121 CCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 122 CCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
Db 181 GTGACGCGGGGACCCCGCGGCTTTCCCGCGGCTGGTGGCCAGTGCCTGTGCTGCTG 240  
QY 182 GTGACGCGGGGACCCCGCGGCTTTCCCGCGGCTGGTGGCCAGTGCCTGTGCTGCTG 241  
Db 241 CCCTGGGACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
QY 242 CCCTGGGACGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301  
Db 301 GAGCTGTGGCGCGAGTGTGCTGAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCG 360  
QY 302 GAGCTGTGGCGCGAGTGTGCTGAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCG 361





QY 2522 CGCATCAGGGCAAGTCTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCTCTCC 2581  
Db 2581 AGCGTGCTGCGAGCTGTGTACGGGACATGGAGACAGCTGTTGGGGGATTCGG 2640  
QY 2582 AGCGTGCTGCGAGCTGTGTACGGGACATGGAGACAGCTGTTGGGGGATTCGG 2641  
Db 2641 CGGACAGGGCTGCTCCTCGGTTGTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACC 2700  
QY 2642 CGGACAGGGCTGCTCCTCGGTTGTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACC 2701  
Db 2701 CACGCGAAACCTTCTCAGGACCTGTGTCAGAGTGTCCTGTAGTAGTGGTGGTGGTG 2760  
QY 2702 CACGCGAAACCTTCTCAGGACCTGTGTCAGAGTGTCCTGTAGTAGTGGTGGTGGTG 2761  
Db 2761 AACTTGGCGAGACAGTGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGACAGGCT 2820  
QY 2762 AACTTGGCGAGACAGTGTGAACCTCCCTGTAGAGACGAGGCCCTGGGTGGACAGGCT 2821  
Db 2821 TTTGTTTCAGATCGCGGCCACGGCCTAT 2848  
QY 2822 TTTGTTTCAGATCGCGGCCACGGCCTAT 2849

RESULT 12  
ID US-08-912-951-3 STANDARD; DNA; UNC; 2176 BP.  
AC xxxxxx

Sequence 3, Application US/08912951  
DE Sequence 3, Application US/08912951  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
CC TITLE OF INVENTION: THERAPEUTIC METHODS  
CC NUMBER OF SEQUENCES: 335  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/912,951  
CC FILING DATE: 14-AUG-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429  
CC REFERENCE/DOCKET NUMBER: 015389-002600US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2176 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: cDNA  
CC FEATURE:  
CC NAME/KEY: -  
CC LOCATION: 1..2176  
CC OTHER INFORMATION: /note= "clone 712562"  
CC SQ SEQUENCE 2176 BP; 432 A; 678 C; 642 G; 422 T; 2 OTHER.

Query Match 45.1%; Score 1711; DB 22; Length 2176;  
Best Local Similarity 91.2%; Pred. No. 0.00e+00;  
Matches 1924; Conservative 0; Mismatches 3; Indels 182; Gaps 1;  
Db 1 GGCACAGTTCCTGCACCTGGCTGATGAGTGTAGTGTGCTGCTGAGTGTCTCAGGTCTTTCTT 60  
QY 1690 GGCACAGTTCCTGCACCTGGCTGATGAGTGTGCTGCTGAGTGTCTCAGGTCTTTCTT 1749  
Db 61 TTATGTACGAGAGACACAGCTTTCAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTG 120  
QY 1750 TTATGTACGAGAGACACAGCTTTCAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTG 1809  
Db 121 GAGCAAGTTGCAAGAGCATTGAATCAGACAGCACTTCAAGAGGGTGCAGCTGCGGGAGCT 180  
QY 1810 GAGCAAGTTGCAAGAGCATTGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT 1869  
Db 181 GTCGGAAGCAGAGGTGACGAGCATTCGGGAAGCAGCGCCGCCCTGTGACGTCCAGACT 240  
QY 1870 GTCGGAAGCAGAGGTGACGAGCATTCGGGAAGCAGCGCCGCCCTGTGACGTCCAGACT 1929  
Db 241 CCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGAGTACGTCTGTGGG 300  
QY 1930 CCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGATTTGTAACATGAGTACGTCTGTGGG 1989  
Db 301 AGCAGAGACGTTCCGACAGAAAGAGCGCGAGCTCTCACCTCGAGGGTGAAGGCACCT 360  
QY 1990 AGCAGAGACGTTCCGACAGAAAGAGCGCGAGCTCTCACCTCGAGGGTGAAGGCACCT 2049  
Db 361 GTTCAGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCTGGCGGCTCTGTGCT 420  
QY 2050 GTTCAGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCTGGCGGCTCTGTGCT 2109  
Db 421 GGGCCTGGACGATATCCACAGGGCCTGGCGACCTTCTGCTGCGTGTGCGGGGCCAGGA 480  
QY 2110 GGGCCTGGACGATATCCACAGGGCCTGGCGACCTTCTGCTGCGTGTGCGGGGCCAGGA 2169  
Db 481 CCGCGCCGCTGAGTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC 540  
QY 2170 CCGCGCCGCTGAGTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC 2229  
Db 541 CCAGGACAGGCTCAGGAGGTTCATCGCCAGCATCATCAAAACCCAGAACACGTACTCGCT 600  
QY 2230 CCAGGACAGGCTCAGGAGGTTCATCGCCAGCATCATCAAAACCCAGAACACGTACTCGCT 2289  
Db 601 GCGTCGATGCGGTGTGTCCAGAGCGCGGCCCATGGGCACGTCCGCAAGGCCCTTCAAGAG 660  
QY 2290 GCGTCGATGCGGTGTGTCCAGAGCGCGGCCCATGGGCACGTCCGCAAGGCCCTTCAAGAG 2349  
Db 661 CCAGTGC----- 667  
|||||



QY 2350 CCACGTCTCTACCTTGACAGACCTCCAGCCGTACATCGGACAGTTCGTGGCTCACCTGCA 2409  
 Db 667 ----- 667  
 QY 2410 GGAGACCAGCCCGCTGAGGGATGCCGTGCTCATCGAGCAGAGCTCCTCCCTGAAATGAGGC 2459  
 Db 667 ----- 667  
 QY 2470 CAGCAGTGGCCCTTTCGACGCTTCTCCATGCTTACGTTGATGTCACACGCGCTGCGCATCAG 2529  
 Db 668 -----CTACGTCAGTCCAGGGGATCCCGCAGGGCTCCCATCTCTCCAGGCTGCT 718  
 QY 2530 GGGCAAGTCTTACGTCAGTCCAGGGATCCCGCAGGGCTCCATCTCTCCAGGCTGCT 2589  
 Db 719 CTGACGCTGTGCTACGCGGACATCGGAGAACAGCTGTTTCGGGGATTCGGCGGGAGCG 778  
 QY 2590 CTGACGCTGTGCTACGCGGACATCGGAGAACAGCTGTTTCGGGGATTCGGCGGGAGCG 2649  
 Db 779 GCTGCTCTGCTGTTGGTGGATTTCTTCTGTTGGTACACCTCACCTCACCCACGCGAA 838  
 QY 2650 GCTGCTCTGCTGTTGGTGGATTTCTTCTGTTGGTACACCTCACCTCACCCACGCGAA 2709  
 Db 839 AACCTTCTCAGGACCTCGTCCGAGGTGTCCTGAGTATGCTGCGTGGTGAACCTTGGC 898  
 QY 2710 AACCTTCTCAGGACCTCGTCCGAGGTGTCCTGAGTATGCTGCGTGGTGAACCTTGGC 2769  
 Db 899 GAAGACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTGTTC 958  
 QY 2770 GAAGACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTGTTC 2829  
 Db 959 GATCGCGCCACAGCCCTATTCCCTGTGTCGGGCTGCTGCTGATACCCCGAGCCCTGGA 1018  
 QY 2830 GATCGCGCCACAGCCCTATTCCCTGTGTCGGGCTGCTGCTGATACCCCGAGCCCTGGA 2889  
 Db 1019 GGTGACAGCAGCTACTCCAGTATGCCCGGACCTCCATCAGAGCAGCTCTCACTTCAA 1078  
 QY 2890 GGTGACAGCAGCTACTCCAGTATGCCCGGACCTCCATCAGAGCAGCTCTCACTTCAA 2949  
 Db 1079 CCGCGGCTTCAAGGCTGGGAGAACATCGCTGCGAAACTCTTTGGGTCCTTTCGGCTGAA 1138  
 QY 2950 CCGCGGCTTCAAGGCTGGGAGAACATCGCTGCGAAACTCTTTGGGTCCTTTCGGCTGAA 3009  
 Db 1139 GTGTCACAGCTGTTCTGGATTTGCAGGTGAACAGCCCTCCAGCGGTGTCACCAACAT 1198  
 QY 3010 GTGTCACAGCTGTTCTGGATTTGCAGGTGAACAGCCCTCCAGCGGTGTCACCAACAT 3069  
 Db 1199 CTACAGATCTCTGCTGCGAGGCTACAGGTTTCCAGCATGTGCTGTCAGCTCCCAT 1258  
 QY 3070 CTACAGATCTCTGCTGCGAGGCTACAGGTTTCCAGCATGTGCTGTCAGCTCCCAT 3129  
 Db 1259 TCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGCGTCTCTCTGACAGCGCTC 1318  
 QY 3130 TCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGCGTCTCTCTGACAGCGCTC 3189  
 Db 1319 CCTGCTACTTCCATCTTGAAGCCAAAGCAGGAGTGTGCTGGGGGCGCAGGGGCG 1378  
 QY 3190 CCTGCTACTTCCATCTTGAAGCCAAAGCAGGAGTGTGCTGGGGGCGCAGGGGCG 3249  
 Db 1379 CGCGGGCCCTGCGCTCCGAGGCGGTGAGTGGCTGTGCCAACAGCATTCCTGCTCAA 1438  
 QY 3250 CGCGGGCCCTGCGCTCCGAGGCGGTGAGTGGCTGTGCCAACAGCATTCCTGCTCAA 3309  
 Db 1439 GCTGACTCGACACCGTGTACCTAGTGTCCACTCTCTGGGGTCACTCAGGACAGCCAGAC 1498  
 QY 3310 GCTGACTCGACACCGTGTACCTAGTGTCCACTCTCTGGGGTCACTCAGGACAGCCAGAC 3369  
 Db 1499 CGAGTGTGAGTGTGAGGCTCCGCGGACAGCGTGTACTGCTGAGGCGCGCAGCCAGCC 1558  
 QY 3370 CGAGTGTGAGTGTGAGGCTCCGCGGACAGCGTGTACTGCTGAGGCGCGCAGCCAGCC 3429  
 Db 1559 GGCAGTGTGCTCAGACTTCAAGACCATCTTGACTGTGGCTGATGGCCCGCCACAGCCAGGC 1618  
 QY 3430 GGCAGTGTGCTCAGACTTCAAGACCATCTTGACTGTGGCTGATGGCCCGCCACAGCCAGGC 3489

RESULT 13

ID US-09-026-981-36 STANDARD; DNA; UNC; 3346 BP.  
 AC xxxxxx

DE Sequence 36, Application US/09026981  
 DE Sequence 36, Application US/09026981  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Counter, Christopher M.  
 CC APPLICANT: Meyerson, Matthew  
 CC APPLICANT: Weinberg, Robert A.  
 CC TITLE OF INVENTION: telomerase Catalytic Subunit Gene and  
 CC NUMBER OF SEQUENCES: 52  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 CC STREET: Two Militia Drive  
 CC CITY: Lexington  
 CC STATE: MA  
 CC COUNTRY: USA  
 CC ZIP: 02173  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patentin Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/09/026,981  
 CC FILING DATE: 20-FEB-1998  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 60/064,322  
 CC FILING DATE: 30-OCT-1997  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 60/055,762  
 CC FILING DATE: 14-AUG-1997  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 60/054,549  
 CC FILING DATE: 01-AUG-1997  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 60/047,151  
 CC FILING DATE: 20-MAY-1997  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 60/038,750  
 CC FILING DATE: 20-FEB-1997  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Granahan, Patricia  
 CC REGISTRATION NUMBER: 32,227  
 CC REFERENCE/DOCKET NUMBER: WHI97-11p4M

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 781-861-6240  
CC TELEFAX: 781-861-9540  
CC INFORMATION FOR SEQ ID NO: 36:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 3346 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
SQ SEQUENCE 3346 BP: 617 A; 1027 C; 965 G; 735 T; 2 OTHER.

Query Match 41.5%; Score 1575; DB 23; Length 3346;  
Best Local Similarity 85.2%; Pred. No. 0.00e+00;  
Matches 1975; Conservative 0; Mismatches 1; Indels 341; Gaps 2;

Db 971 GGGTTGGCTGTTCCTCCGGCCGAGACACCGTCTCGTGAGGAGATCCTGGCCAAAGTTCC 1030  
Qy 1641 GGGTTGGCTGTTCCTCCGGCCGAGACACCGTCTCGTGAGGAGATCCTGGCCAAAGTTCC 1700  
Db 1031 TGCACCTGGCTGATGATGTGTACGTGTCGAGCTGCTCAGGTCTCTTTTATGTCACGG 1090  
Qy 1701 TGCACCTGGCTGATGATGTGTACGTGTCGAGCTGCTCAGGTCTCTTTTATGTCACGG 1760  
Db 1091 AGACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGACGAGTTGC 1150  
Qy 1761 AGACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGACGAGTTGC 1820  
Db 1151 AAAGCATTGGATCAGACGACTTGAAGAGGGTGCAGCTCGGGAGGTGTCGGAAGCAG 1210  
Qy 1821 AAAGCATTGGATCAGACGACTTGAAGAGGGTGCAGCTCGGGAGGTGTCGGAAGCAG 1880  
Db 1211 AGGTGAGCAGCATCGGAAGCAGCGCCGCTGCTGACGTCCAGATCTCGCTTCATCC 1270  
Qy 1881 AGGTGAGCAGCATCGGAAGCAGCGCCGCTGCTGACGTCCAGATCTCGCTTCATCC 1940  
Db 1271 CCAAGCCTGACGGCTGCGCGGATGTGAACATGACTAGTCTGTGGAGCCAGAGCT 1330  
Qy 1941 CCAAGCCTGACGGCTGCGCGGATGTGAACATGACTAGTCTGTGGAGCCAGAGCT 2000  
Db 1331 TCCGCAGAGAAAGAGGCGCAGCGCTCACCTCGAGGGTGAAGCACTGTTACGCGTGC 1390  
Qy 2001 TCCGCAGAGAAAGAGGCGCAGCGCTCACCTCGAGGGTGAAGCACTGTTACGCGTGC 2060  
Db 1391 TCAACTACGAGGCGCGGCGCCCGCTGCTGGGCGCTCTGTGTGGGCTGTGAGC 1450  
Qy 2061 TCAACTACGAGGCGCGGCGCCCGCTGCTGGGCGCTCTGTGTGGGCTGTGAGC 2120  
Db 1451 ATATCCACAGGCGCTGCGCAGCTTCTGCTGCTGCTGCGGCGCCAGACCGCGCGCTG 1510  
Qy 2121 ATATCCACAGGCGCTGCGCAGCTTCTGCTGCTGCTGCGGCGCCAGACCGCGCGCTG 2180  
Db 1511 AGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGACACCATCCCCAGGACAGGC 1570  
Qy 2181 AGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGACACCATCCCCAGGACAGGC 2240  
Db 1571 TCACGGAGGTCTATCCAGCATCATCAAAACCCAGAACACGTACGTGCGTGGGTATG 1630  
Qy 2241 TCACGGAGGTCTATCCAGCATCATCAAAACCCAGAACACGTACGTGCGTGGGTATG 2300  
Db 1631 CGGTGCTCCAGAGCCCGCCATGGCAGCTCCGCAAGCGCTTCAAGAGCCACGTC 1686  
Qy 2301 CGGTGCTCCAGAGCCCGCCATGGCAGCTCCGCAAGCGCTTCAAGAGCCACGTCCTA 2360  
Db 1686 ----- 1686  
Qy 2361 CTTTGACAGACCTCCAGCGGTACATGCGACAGTTGCTGCTCACCTGCGAGACACGCC 2420  
Db 1686 ----- 1686  
Qy 2421 CGCTGAGGGATGCCGTCTCATTCGACAGAGCTCTCCCTGAATGAGGCCAGCAGTGCC 2480  
Db 1687 -----CT 1688  
||

Qy 2481 TCCTCGAGCTTCTCTAGGCTTCAATGTGCCACACGCGGTGGCGATCAGGGGCAAGTCCT 2540  
Db 1689 ACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATTCCTCTCCACGCTCTCTGACGCTGT 1748  
Qy 2541 ACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATTCCTCTCCACGCTCTCTGACGCTGT 2600  
Db 1749 GCTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGCTGCTCTGC 1808  
Qy 2601 GCTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGCTGCTCTGC 2660  
Db 1809 GTTTGGTGGATGATTCTTGTGTGTGACACTCACCTCACCCAGCGCAAAACCTTCCTCA 1868  
Qy 2661 GTTTGGTGGATGATTCTTGTGTGTGACACTCACCTCACCCAGCGCAAAACCTTCCTCA 2720  
Db 1869 GGACCCCTGGTCCGAGGTGTCCTGAGTATGCTGCTGGTGAACCTTTCGGAAGACACTGG 1928  
Qy 2721 GGACCCCTGGTCCGAGGTGTCCTGAGTATGCTGCTGGTGAACCTTTCGGAAGACACTGG 2780  
Db 1929 TGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGATGCCGCCC 1988  
Qy 2781 TGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGATGCCGCCC 2840  
Db 1989 ACGSCCTATTCCCTGCTGCGGCTGCTGCTGGATACCCGAGCCCTCGAGGTGCAGAGCG 2048  
Qy 2841 ACGSCCTATTCCCTGCTGCGGCTGCTGCTGGATACCCGAGCCCTCGAGGTGCAGAGCG 2900  
Db 2049 ACTACTCCAGTATGCCCGGACCTTCCATCAGAGCCAGCTCTCACCTTCAACCGCGGCTTCA 2108  
Qy 2901 ACTACTCCAGTATGCCCGGACCTTCCATCAGAGCCAGCTCTCACCTTCAACCGCGGCTTCA 2960  
Db 2109 AGGTGAGGAGAACATCGCTCGCAAACTCTTTGGGGTCTTGGGGCTGAAAGTGTACAGCC 2168  
Qy 2961 AGGTGAGGAGAACATCGCTCGCAAACTCTTTGGGGTCTTGGGGCTGAAAGTGTACAGCC 3020  
Db 2169 TGTTCCTGGATTTGCAAGTGAACAGCCCTCCAGAGGTGTCACCAACATCTACAGATCC 2228  
Qy 3021 TGTTCCTGGATTTGCAAGTGAACAGCCCTCCAGAGGTGTCACCAACATCTACAGATCC 3080  
Db 2229 TCCTGCTGCGAGCGTACAGGTTTTCACGATGTGTGCTGCGAGCTCCCATTTTCATCAGCAAG 2288  
Qy 3081 TCCTGCTGCGAGCGTACAGGTTTTCACGATGTGTGCTGCGAGCTCCCATTTTCATCAGCAAG 3140  
Db 2289 TTGGGAAGACCCACATTTTTCCTGCGCTCATCTCTGACAGGCTCTCCTCTGCTACT 2348  
Qy 3141 TTGGGAAGACCCACATTTTTCCTGCGCTCATCTCTGACAGGCTCTCCTCTGCTACT 3200  
Db 2349 CCATCTGGAAGCCAAAGACGACGCGGAGAAACATTTCTGTGCTGACTCCTCGGCTGC 2408  
Qy 3201 CCATCTGGAAGCCAAAGACGCG - - - - - 3224  
Db 2409 TTGGGTCCGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTCTGGGTGTGGGCAAGCTTC 2468  
Qy 3224 ----- 3224  
Db 2469 CGGTGCTCTCTGGAGGGGAGCTGGGGCTGGGCTGTGACTCTCTCAGCTCTGTTTCCCC 2528  
Qy 3224 ----- 3224  
Db 2529 CAGGGATGTCGTGGGGCCAAAGGGCGCCCGCGCCCTCTGCGCTCCGAGGCCGTGCAGT 2588  
Qy 3225 ---GGATGTGCTGGGGCCAAAGGGCGCCCGCGCCCTCTGCGCTCCGAGGCCGTGCAGT 3281  
Db 2589 GGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTTCACACCGCTGTCCACTACGTGCAC 2648  
Qy 3282 GGCTGTGCCACCAAGCATTCCTGCTCAAGTGTACTTCACACCGCTGTCCACTACGTGCAC 3341  
Db 2649 TCCTGGGGTCACTCAGGACAGCCCGCAGAGCTGAGTTCGGAAGCTCCCGGGGACGAGC 2708  
Qy 3342 TCCTGGGGTCACTCAGGACAGCCCGCAGAGCTGAGTTCGGAAGCTCCCGGGGACGAGC 3401  
Db 2709 TGACTGCCCTGGAGGCCGAGCCAAACCGCACTGCCCTCAGACTTCAAGACCACTCTGG 2768  
Qy 3402 TGACTGCCCTGGAGGCCGAGCCAAACCGCACTGCCCTCAGACTTCAAGACCACTCTGG 3461



QY 2411 GAGACACCGCGCTGAGGATGCCGTCTCATCAGCAGAGCTCCTCCTGAATGAGCC 2470  
Db 665 ----- 665  
QY 2471 AGCAGTGGCCTCTTTCGACGCTTCTCCTAGCTTTCATGTGCCACACGCGCTGGCATCAGG 2530  
Db 666 -----CTACGTCACAGTGCACGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTC 717  
QY 2531 GGCAAGTCTCCTACGTCACAGTGCACGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTC 2590  
Db 718 TGACGCTGTGCTACGGCGACATGGAACAAGCTGTTTGGGGGATTCGGCGGACGGG 777  
QY 2591 TGACGCTGTGCTACGGCGACATGGAACAAGCTGTTTGGGGGATTCGGCGGACGGG 2650  
Db 778 CTGCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAGCGGAAA 837  
QY 2651 CTGCTCTCGGTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAGCGGAAA 2710  
Db 838 ACCTTCCTCAGGACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGAACCTTCGG 897  
QY 2711 ACCTTCCTCAGGACCTGTGTCGAGGTGTCCCTGAGTATGGCTGCGTGAACCTTCGG 2770  
Db 898 AAGACAGTGTGAATCTCCCTGTAGAAGACGAGCCCTGGGTGGCAGCGCTTTTGTTCAG 957  
QY 2771 AAGACAGTGTGAATCTCCCTGTAGAAGACGAGCCCTGGGTGGCAGCGCTTTTGTTCAG 2830  
Db 958 ATGCGCGCCACGCGCTATTCCCTGTGTGGCGCTGCTGTGATACCCGACCCCTGGAG 1017  
QY 2831 ATGCGCGCCACGCGCTATTCCCTGTGTGGCGCTGCTGTGATACCCGACCCCTGGAG 2890  
Db 1018 GTGACAGCGACTACTCCAGTATGCCCGGACCTCCATCAGAGCAGTCTCACCTTCAAC 1077  
QY 2891 GTGACAGCGACTACTCCAGTATGCCCGGACCTCCATCAGAGCAGTCTCACCTTCAAC 2950  
Db 1078 CGCGGCTTAAAGGCTGGGAGAACATCGTCGCAAACTCTTTGGGGTCTTGGCGCTGAAG 1137  
QY 2951 CGCGGCTTAAAGGCTGGGAGAACATCGTCGCAAACTCTTTGGGGTCTTGGCGCTGAAG 3010  
Db 1138 TGTACAGCGCTGTTTCTGGAATTTGAGTTTGCAGGTGAACAGCCTCCAGAGGTTGTCACCAACATC 1197  
QY 3011 TGTACAGCGCTGTTTCTGGAATTTGAGTTTGCAGGTGAACAGCCTCCAGAGGTTGTCACCAACATC 3070  
Db 1198 TACAAGATCTCTGCTGCAGGCGTACAGTTTCACGCATGTGTGCTGAGCTCCCATTT 1257  
QY 3071 TACAAGATCTCTGCTGCAGGCGTACAGTTTCACGCATGTGTGCTGAGCTCCCATTT 3130  
Db 1258 CATCAGCAAGTTTGGAAAGAACCC-ACATTTTCTCGCGCTCATCTCTGACAGGC-TCC 1315  
QY 3131 CATCAGCAAGTTTGGAAAGAACCCACATTTTCTCGCGCTCATCTCTGACAGGCCTCC 3190  
Db 1316 CTCTGCTACTCCATCTCTGAAAGCAAGAACGACGGATGCTGTGGGGGCCAAGGGCGCC 1375  
QY 3191 CTCTGCTACTCCATCTCTGAAAGCAAGAACGACGGATGCTGTGGGGGCCAAGGGCGCC 3250  
Db 1376 GCCGGCCCTCTGCCCTCCGAGGCGTGCAGTGGTGTGCCACCAAGCATTCCTGCTCAAG 1435  
QY 3251 GCCGGCCCTCTGCCCTCCGAGGCGTGCAGTGGTGTGCCACCAAGCATTCCTGCTCAAG 3310  
Db 1436 CTGACTCGACACCGTGTCACTAGTGCCTCCTCGGGTCACTCAGGACAGCCAGACG 1495  
QY 3311 CTGACTCGACACCGTGTCACTAGTGCCTCCTCGGGTCACTCAGGACAGCCAGACG 3370  
Db 1496 CAGCTGAGTCGGAAGCTCCCGGGACACGCTGACTGCCCTGGAGGCGCGAGCAACCCG 1555  
QY 3371 CAGCTGAGTCGGAAGCTCCCGGGACACGCTGACTGCCCTGGAGGCGCGAGCAACCCG 3430  
Db 1556 GCAGTGCCTTCAGACTTCAAGACCATCCTTGAGTATGGCCACCCGCGCCACAGCGGCC 1615  
QY 3431 GCAGTGCCTTCAGACTTCAAGACCATCCTTGAGTATGGCCACCCGCGCCACAGCGGCC 3490  
Db 1616 GAGACGACACACGAGCGCTGTGACGCGGGCTTATAGTCTCCAGGGAGGGGGCG 1675  
QY 3491 GAGACGACACACGAGCGCTGTGACGCGGGCTC-TACGTCCCGAGGGAGGGGGCG 3549

Db 1676 GCCACACCCAGCGCTGCACCGCTGGAGTCTGAGGCTTGAAGTGTGTTGGCCGAGGC 1735  
QY 3550 GCCACACCCAGCGCGCCGCTGGAGTCTGAGGCTTGAAGTGTGTTGGCCGAGGC 3609  
Db 1736 CTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGACAGCAAGG 1795  
QY 3610 CTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCTGAGGCTGACAGCAAGG 3669  
Db 1796 GCTGAGTGTCCAGCACACCTGC-GTTTTTCACTTCCACAGGCTGGCTTCG-TCCACC 1853  
QY 3670 GCTGAGTGTCCAGCACACCTGCGCTTCACTTCCACAGGCTGGCTTCGCTCCACC 3729  
Db 1854 CCAGGCGCCAGCTTTTCTCACCAGGAGCGGCTTCACTTCCACATAGGAATAGTCCA 1913  
QY 3730 CCAGGCGCCAGCTTTTCTCACCAGGAGCGGCTTCACTTCCACATAGGAATAGTCCA 3789  
Db 1914 TCCCCAGAT 1922  
QY 3790 TCCCCGTAT 3798  
RESULT 15  
ID US-08-954-050-100 STANDARD; DNA; UNC; 2171 BP.  
AC xxxxxx  
DT  
Sequence 100, Application US/08854050  
Sequence 100, Application US/08854050  
CC GENERAL INFORMATION:  
CC APPLICANT: Cech, Thomas R.  
CC APPLICANT: Lingner, Joachim  
CC APPLICANT: Nakamura, Toru  
CC APPLICANT: Chapman, Karen B.  
CC APPLICANT: Morin, Gregg B.  
CC APPLICANT: Harley, Calvin  
CC APPLICANT: Andrews, William H.  
CC TITLE OF INVENTION: Novel Telomerase  
CC NUMBER OF SEQUENCES: 225  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Townsend and Townsend and Crew LLP  
CC STREET: Two Embarcadero Center, 8th Floor  
CC CITY: San Francisco  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 94111  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/854,050  
CC FILING DATE: 09-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/851,843  
CC FILING DATE: 06-MAY-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/846,017  
CC FILING DATE: 25-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/844,419  
CC FILING DATE: 18-APR-1997  
CC CLASSIFICATION: 536  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/724,643  
CC FILING DATE: 01-OCT-1996  
CC CLASSIFICATION: 536  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Apple, Randolph T.  
CC REGISTRATION NUMBER: 36,429

CC REFERENCE/DOCKET NUMBER: 015389-002930US  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 576-0200  
CC TELEFAX: (415) 576-0300  
CC INFORMATION FOR SEQ ID NO: 100:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 2171 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC FEATURE:  
CC NAME/KEY: CDS  
CC LOCATION: 22..1716

CC OTHER INFORMATION: /note= "EcoRI-NotI insert of  
CC OTHER INFORMATION: clone 712562 encoding 63 kDa  
CC OTHER INFORMATION: telomerase protein"  
CC SEQUENCE 2171 BP; 433 A; 667 C; 641 G; 429 T; 1 OTHER.

Query Match 40.0%; Score 1520; DB 21; Length 2171;  
Best Local Similarity 90.8%; Pred. No. 0.00e+00;  
Matches 1915; Conservative 1; Mismatches 5; Indels 188; Gaps 7;  
Db 1 GCCAAGTTCCTGCACTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGTCTTTCTTT 60  
Qy 1691 GCCAAGTTCCTGCACTGCTGATGAGTGTACGTCGTCGAGCTGCTCAGTCTTTCTTT 1750  
Db 61 TATGTCAGGAGACACACCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTCTG 120  
Qy 1751 TATGTCAGGAGACACACCTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTCTG 1810  
Db 121 AGCAAGTTGCAAGCATTTGGAATCAGACAGACACTTTGAAGAGGTCAGCTGCGGGAGCTG 180  
Qy 1811 AGCAAGTTGCAAGCATTTGGAATCAGACAGACACTTTGAAGAGGTCAGCTGCGGGAGCTG 1870  
Db 181 TCGAAGCAGAGGTCAGCAGCATCGGAGCCAGCCGCGCTGCTCAGCTCCAGACTC 240  
Qy 1871 TCGAAGCAGAGGTCAGCAGCATCGGAGCCAGCCGCGCTGCTCAGCTCCAGACTC 1930  
Db 241 CGCTTCATCCCAAGCCTGACGGGCTGGGGCGGATTTGAACATGGACTAGCTCTGGGA 300  
Qy 1931 CGCTTCATCCCAAGCCTGACGGGCTGGGGCGGATTTGAACATGGACTAGCTCTGGGA 1990  
Db 301 GCCAAGAGTTCCCGCAGAGAAAGAGCCGAGCGCTCTCACTCAGAGGTTGAAGCACTG 360  
Qy 1991 GCCAAGAGTTCCCGCAGAGAAAGAGCCGAGCGCTCTCACTCAGAGGTTGAAGCACTG 2050  
Db 361 TTCAGCGTGTCAACTACGAGCGGCGCGCGCGCCCGCTCCTGGGCGCTCTGTGCTG 420  
Qy 2051 TTCAGCGTGTCAACTACGAGCGGCGCGCGCGCGCCCGCTCCTGGGCGCTCTGTGCTG 2110  
Db 421 GGCCTGGAGCATATCCAGAGGCTGGCGCACCTTCGTGCTGCTGCTGGGCGCCAGGAC 480  
Qy 2111 GGCCTGGAGCATATCCAGAGGCTGGCGCACCTTCGTGCTGCTGCTGGGCGCCAGGAC 2170  
Db 481 CGCGCGCTGAGCTGCTACTTTGCAAGTGTGATGTGACGGGCGGTACGACACCATCCCC 540  
Qy 2171 CGCGCGCTGAGCTGCTACTTTGCAAGTGTGATGTGACGGGCGGTACGACACCATCCCC 2230  
Db 541 CAGGACAGGCTCAGGAGGTATCCGAGCATCATCAACCCAGAAACAGTACTGCGTG 600  
Qy 2231 CAGGACAGGCTCAGGAGGTATCCGAGCATCATCAACCCAGAAACAGTACTGCGTG 2290  
Db 601 COTCGGTATCGGTGTCAGAAAGCCGCC - ATGGGACGCTCCGCAAGGCTTCAAGAGC 659  
Qy 2291 COTCGGTATCGGTGTCAGAAAGCCGCCCATGGGACGTCCTGCAAGGCTTCAAGAGC 2350  
Db 660 CAGGTC----- 665  
Qy 2351 CAGGTCCTTACCTTGACAGACCTCCAGCGGTACATCGGACATTCGTGGCTCAGCTGCGAG 2410  
Db 665 ----- 665

Qy 2411 GAGACCAGCCGCTGAGGGATGCCGTGTCATCGACGAGAGCTCCTCCTGAAATGAGGCC 2470  
Db 665 ----- 665  
Qy 2471 AGCAGTGGCCTCTTCGACGCTCTTCTTACGCTTTCATGTGCCACACGCGGTGCGCATCAGG 2530  
Db 666 -----CTACGTCAGTGGCAGGGATCCGCGAGGGTCCATCTCTCCACGCTGCTC 717  
Qy 2531 GGCAAGTCTTACGTCAGTGGCAGGGATCCGCGAGGGTCCATCTCTCCACGCTGCTC 2590  
Db 718 TGCAGCTGTGTCTACGCGCATGAGAAAGCTTTTGGGGGATTCGGGCGGACGGG 777  
Qy 2591 TGCAGCTGTGTCTACGCGCATGAGAAAGCTTTTGGGGGATTCGGGCGGACGGG 2650  
Db 778 CTGCTCTCGGTTTGGTGGATGATTTCTTTGTTGTGACACCTCACCACCGCGAAA 837  
Qy 2651 CTGCTCTCGGTTTGGTGGATGATTTCTTTGTTGTGACACCTCACCACCGCGAAA 2710  
Db 838 ACCTTCTCAGGACCTTGGTCCGAGGTGCTCCTGATATGGCTCGGTGGTGAACCTTGGGG 897  
Qy 2711 ACCTTCTCAGGACCTTGGTCCGAGGTGCTCCTGATATGGCTCGGTGGTGAACCTTGGGG 2770  
Db 898 AAGACAGTGTGAACTTCCCTGTAGAAAGAGAGGCGCTGGTGGCAGGCTTTTGTTCAG 957  
Qy 2771 AAGACAGTGTGAACTTCCCTGTAGAAAGAGAGGCGCTGGTGGCAGGCTTTTGTTCAG 2830  
Db 958 ATGCGCGCCACGCGCTATTCCCTGGTGGGCGCTGCTGTGGATACCGGACCTTGAG 1017  
Qy 2831 ATGCGCGCCACGCGCTATTCCCTGGTGGGCGCTGCTGTGGATACCGGACCTTGAG 2890  
Db 1018 GTGACAGAGCTACTCCAGCTATGCGCGGACCTCCATCAGAGCGAGTCTCACCCTTCAAC 1077  
Qy 2891 GTGACAGAGCTACTCCAGCTATGCGCGGACCTCCATCAGAGCGAGTCTCACCCTTCAAC 2950  
Db 1078 CGCGGCTTCAAGGCTGGGAGAACATGCGTGCAGAACTCTTTGGGGTCTTCGCGCTGAAG 1137  
Qy 2951 CGCGGCTTCAAGGCTGGGAGAACATGCGTGCAGAACTCTTTGGGGTCTTCGCGCTGAAG 3010  
Db 1138 TGTACAGAGCTGTTCTGGAATTTGACAGTGAACAGCTCCAGAGCGGTGTGCACCAACATC 1197  
Qy 3011 TGTACAGAGCTGTTCTTGGATTTGACAGTGAACAGCTCCAGAGCGGTGTGCACCAACATC 3070  
Db 1198 TACAAGATCTCTGCTGACAGCGCTACAGTTTCAACGATGTGTGCTGACAGTCCCATTT 1257  
Qy 3071 TACAAGATCTCTGCTGACAGCGCTACAGTTTCAACGATGTGTGCTGACAGTCCCATTT 3130  
Db 1258 CATCAGAAAGTTTGAAGAACCC - ACATTTTCTGCGCTCATCTCTGACACAGGC - TCC 1315  
Qy 3131 CATCAGAAAGTTTGAAGAACCCCATTTTCTGCGCGTCACTCTGACACAGCGCTCC 3190  
Db 1316 CTCTGCTACTCATCTCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCCAAGGGCGCC 1375  
Qy 3191 CTCTGCTACTCATCTCTGAAAGCCAAAGACGAGGATGTGCTGGGGCCCAAGGGCGCC 3250  
Db 1376 GCCGGCCTCTGCGCTCCGAGGCGGTGACGTGTGTGCCACCAAGCATTCCTGCTCAAG 1435  
Qy 3251 GCCGGCCTCTGCGCTCCGAGGCGGTGACGTGTGTGCCACCAAGCATTCCTGCTCAAG 3310  
Db 1436 CTGACTGACACCGTGTACCTACGTGCGACTCTCTGGGTCACTCAGACAGCCAGAGC 1495  
Qy 3311 CTGACTGACACCGTGTACCTACGTGCGACTCTCTGGGTCACTCAGACAGCCAGAGC 3370  
Db 1496 CAGCTGAGTCGGAAGCTCCCGGGGACGAGCTGACTGCCCTGGAGGCGGAGCCCAACCCG 1555  
Qy 3371 CAGCTGAGTCGGAAGCTCCCGGGGACGAGCTGACTGCCCTGGAGGCGGAGCCCAACCCG 3430  
Db 1556 GCATGCGCTCAGACTTCAAGACCATCTGAGCTGATGGCCACCGCCGCGCCACAGCGAGCC 1615  
Qy 3431 GCATGCGCTCAGACTTCAAGACCATCTGAGCTGATGGCCACCGCCGCGCCACAGCGAGCC 3490  
Db 1616 GAGAGCAGACACAGAGCGCTGTACGCGCGGCTTTATGCTCCAGGAGGAGGGGCG 1675  
Qy 3491 GAGAGCAGACACAGAGCGCTGTACGCGCGGCTGTACGCGCGGCTC - TACGTCCAGGGAGGGGCG 3549

```

Db 1676 GCCCACCCAGGCTGCACCGCTGGAGTCTGAGGCCCTGAGTGTGTTGGCCGAGGC 1735
QY 3550 GCCCACCCAGGCGCGCACCGCTGGAGTCTGAGGCCCTGAGTGTGTTGGCCGAGGC 3609
Db 1736 CTGCATGTCGGCTGAAGCTGAGTGTCCGGCTGAGGCTGAGCGAGTGTCCAGCCAGG 1795
QY 3610 CTGCATGTCGGCTGAAGCTGAGTGTCCGGCTGAGGCTGAGCGAGTGTCCAGCCAGG 3669
Db 1796 GGTGAGTGTCCAGCACACCTGC-GTTTTCACTTCCCACAGGCTGGCGTTCCG-TCCACC 1853
QY 3670 GGTGAGTGTCCAGCACACCTGCGGTCTTCACTTCCCACAGGCTGGCGTCCGCTCCACC 3729
Db 1854 CCAGGCCAGCTTTTCTCACCAGGAGCCGGCTTCCACTCCCCACATAGGAATAGTCCA 1913
QY 3730 CCAGGCCAGCTTTTCTCACCAGGAGCCGGCTTCCACTCCCCACATAGGAATAGTCCA 3789
Db 1914 TCCCCAGAT 1922
QY 3790 TCCCCTGAT 3798

```

Search completed: Fri Dec 25 08:54:06 1998  
Job time : 8861 secs.